Evaluation of the Ozone Fields in NASA's MERRA-2 Reanalysis

Krzysztof Wargan[1,2], Gordon Labow[2,3], Stacey Frith[2,3], Steven Pawson[1], Nathaniel Livesey[4] and Gary Partyka[1,2]

[1] Global Modeling and Assimilation Office, Code 610.1, NASA Goddard Space Flight Center, Greenbelt, MD

[2] Science Systems and Applications Inc., Lanham, MD

[3] Atmospheric Chemistry and Dynamics Laboratory, Code 614, NASA Goddard Space Flight Center, Greenbelt, MD

[4] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA

Correspondence: Krzysztof Wargan, email: Krzysztof.Wargan-1@nasa.gov

Abstract. We describe and assess the quality of the assimilated ozone product from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) produced at NASA's Global Modeling and Assimilation Office (GMAO) spanning the time period from 1980 to present. MERRA-2 assimilates partial column ozone retrievals from a series of Solar Backscatter Ultraviolet (SBUV) radiometers on NASA and NOAA spacecraft between January 1980 and September 2004; starting in October 2004 retrieved ozone profiles from the Microwave Limb Sounder (MLS) and total column ozone from the Ozone Monitoring Instrument on NASA's EOS Aura satellite are assimilated. We compare the MERRA-2 ozone with independent satellite and ozonesonde data focusing on the representation of the spatial and temporal variability of stratospheric and upper tropospheric ozone and on implications of the change in the observing system from SBUV to EOS Aura. The comparisons show agreement within 10 % (standard deviation of the difference) between MERRA-2 profiles and independent satellite data in most of the stratosphere. The agreement improves after 2004 when EOS Aura data are assimilated. The standard deviation of the differences between the lower stratospheric and upper tropospheric MERRA-2 ozone and ozonesondes is 11.2 % and 24.5 %, respectively, with correlations of 0.8 and above, indicative of a realistic representation of the near-tropopause ozone variability in MERRA-2. The agreement improves significantly in the EOS Aura period, however MERRA-2 is biased low in the upper troposphere with respect to the ozonesondes. Caution is recommended when using MERRA-2 ozone for decadal changes and trend studies.

1. Introduction

Atmospheric reanalyses produce global high spatial and temporal resolution long-term records of meteorological fields and composition of earth's atmosphere by utilizing the data assimilation methodology (Cohn 1997; Kalnay 2003), whereby satellite and ground-based observations are combined with general circulation model (GCM) simulations in a statistically optimal way. The Modern Era Retrospective Analysis for Research and Applications (MERRA: Rienecker et al. 2011) was the first reanalysis generated using the Goddard Earth Observing System (GEOS) Data Assimilation System (DAS) by NASA's Global Modeling and Assimilation Office (GMAO). MERRA, first released in 2009, covered the years 1979-2015 (production ended on 29 February 2016). It was followed by the recently released MERRA-2 dataset (Bosilovich et al. 2015), which is the focus of this paper.

While most reanalyses include assimilated ozone fields, the overall lack of validation and uncertain quality of these fields has not encouraged the atmospheric ozone community to use them in scientific research. Typically, researchers prefer to utilize satellite and in-situ ozone data along with assimilated meteorological variables. To our knowledge, the only comprehensively validated reanalysis ozone fields are those from the European Centre for Medium-Range Weather Forecasts reanalyses: ERA-40 (Dethof and Hólm 2004) and ERA-Interim (Dragani 2011). On the other hand, a large body of literature evaluates ozone data assimilation results from (usually relatively short) assimilation experiments (see Lahoz et al. 2014). In addition, significant effort has been invested in the evaluation of multiyear

67   chemical analyses under the Monitoring Atmospheric Composition and Climate (MACC)

68   project conducted using the ECMWF Integrated Forecast system coupled with

69   comprehensive chemistry models (e.g., Inness et al., 2013; Inness et al.; 2015l; Lefever et

70   al. 2015). Much of this work has demonstrated an added value brought to satellite

71   observations of ozone through data assimilation. In particular, work done at the GMAO over

72   the past decade has shown that assimilation of retrieved ozone data from the Microwave

73   Limb Sounder (MLS) along with total ozone observations from the Ozone Monitoring

74   Instrument (OMI), both onboard the Earth Observing System (EOS) Aura satellite, produces

75   realistic global distributions of ozone in the stratosphere and upper troposphere (Stajner et

76   al. 2008; Wargan et al. 2015). Ziemke et al. (2014) compared tropospheric ozone derived

77   from the GEOS DAS with that from trajectory mapping of the Earth Observing System

78   (EOS) Aura data and direct profile retrieval from OMI radiances and concluded that data

79   assimilation is the best of these three strategies at generating global ozone product. By

80   combining available measurements with global circulation model short-term forecasts the

81   data assimilation methodology allows the propagation of observational information by

82   assimilated winds resulting in global 3-dimensional maps of ozone concentrations at spatial

83   and temporal resolutions far exceeding those attainable with satellite data alone. For

84   example, a satellite-borne instrument at a sun-synchronous orbit typically samples about 30

85   points along a latitude circle per day (15 if day-only observations are made), corresponding

86   to a 12° resolution in the longitudinal direction. By contrast, the resolution of MERRA-2 is

87   approximately 0.625° in longitude and has a three-hourly output frequency.

88

89　This paper presents a description and evaluation of the MERRA-2 ozone product against

90　independent satellite and ozonesonde measurements. It is intended to provide guidance for

91　researchers who may wish to use the MERRA-2 ozone product in scientific studies. We

92　focus on the following questions:

93

94　1.　How well does MERRA-2 represent the spatio-temporal variability of the

95　　　stratospheric and upper-tropospheric ozone?

96　2.　What are the impacts of changing the ozone observing system in MERRA-2 in late

97　　　2004 when new types of data were introduced?

98

99　The first question addresses the ability of the assimilated ozone to represent the statistical

100　characteristics of the actual ozone field, which must be established before the assimilated

101　fields can be used in research studies. In particular we focus on results in the lower

102　stratospheric/upper tropospheric region, where variability is enhanced and high spatio-

103　temporal resolution satellite observations are limited. Analysis of, for example, radiative

104　forcing near the tropopause, or stratosphere-troposphere exchange mechanisms require both

105　consistency between the constituent and dynamical fields, and accurate representation of

106　small-scale variability. The stability of the multi-decadal MERRA-2 ozone record and the

107　evolution of the product's quality resulting from changes in the observing system are

108　addressed in the second question.

109

110　While this study is not intended to be a comprehensive validation of the original MERRA

111　reanalysis we include both MERRA and MERRA-2 ozone in a number of comparisons in

112   order to emphasize the differences that arise (mainly) from differences in the input ozone

113   observations between the two products.

114

115   The paper is organized as follows. Section 2 describes the MERRA-2 reanalysis, focusing

116   on the treatment of ozone. The MERRA-2 ozone observing system is discussed in Section 3.

117   A description of independent data used for validation is given in Section 4. The differences

118   between MERRA and MERRA-2 relevant to ozone are summarized in Section 5. Section 6

119   presents the results of comparisons of the MERRA-2 ozone fields against satellite and

120   ozonesonde observations. Section 7 summarizes the results of this study.

121

122   **2. The MERRA-2 reanalysis**

123

124   MERRA-2 (Bosilovich et al. 2015) is a multi-year reanalysis developed at GMAO, covering

125   the 'satellite era' of earth observations from 1980 to present. It is produced using the

126   Version 5.12.4 of the GEOS DAS. Gridded data are released at a 0.625° by 0.5° longitude

127   by latitude resolution at 72 sigma-pressure hybrid layers between the surface and 0.01 hPa.

128   The bottom 32 layers are terrain-following while remaining model layers, from 164 to .01

129   hPa, are constant pressure surfaces. MERRA-2 replaces the original GMAO reanalysis,

130   MERRA (Rienecker et al. 2011). Both the general circulation model and the observing

131   system were significantly updated since MERRA. The model updates include a transition

132   from a regular spherical grid to the cubed sphere (Putman and Lin 2007). A re-tuning of the

133   gravity wave parameterization affects the stratosphere, in particular resulting in a realistic

134   quasi-biennial oscillation (QBO) generated by the model in the absence of observations

135   (Coy et al. 2016). Other changes include upgrades to moist physics and the enforcement of

136   dry mass conservation in the assimilation (Takacs et al. 2016). The model updates are

137   described in detail in Molod et al. (2015). The observing system was expanded to include

138   more recent satellite data: the Infrared Atmospheric Sounding Interferometer (IASI, starting

139   in September 2008), the Cross-Track Infrared Sounder (on the Suomi-NPP satellite, from

140   April 2012 onward) and Advanced Technology Microwave Sounder (on Suomi-NPP,

141   starting in November 2011) in addition to those already used in MERRA. Radiance data

142   from the Stratospheric Sounding Unit instruments are used in both reanalyses but in

143   MERRA-2 these observations are assimilated with the more advanced Community

144   Radiative Transfer Model, the same as for all other radiance data. Both reanalyses assimilate

145   radiance observations from the Advanced Microwave Sounding Unit and Atmospheric

146   Infrared Sounder. A comprehensive description of the MERRA-2 observing system is given

147   in McCarty et al. (2016). An adaptive bias correction scheme is applied to all radiance data

148   and some aircraft temperature observations (McCarty et al. 2016). This study uses the

149   three-hourly MERRA-2 assimilated ozone output on the native 72 levels (GMAO 2015a)

150   and the one-hourly total ozone product (GMAO 2015b).

151

152

153   **3. Assimilated ozone data**

154   The ozone data sources used in MERRA-2 are listed in Table 1. Between January 1980 and

155   September 2004 partial column and total ozone from a series of Solar Backscatter

156   Ultraviolet (SBUV) instruments on NASA and NOAA spacecraft are assimilated. Starting 1

157   October 2004 these observations are turned off and replaced by total column ozone data

158 from OMI and stratospheric ozone profiles from MLS, both on NASA's EOS Aura satellite.

159 The MERRA-2 ozone record is thus divided into two periods, henceforth referred to as the

160 SBUV period (January 1980- September 2004) and the Aura period (from October 2004

161 onward). Joint assimilation of MLS and SBUV data was considered but tests showed that it

162 would generate unphysical features in the assimilated ozone fields in the tropical and

163 subtropical stratosphere. Since the GEOS DAS currently ingests SBUV partial columns

164 without accounting for averaging kernels, an apparent bias between MLS and SBUV arises

165 leading to large analysis increments that produce satellite track-following features in the

166 ozone fields.

167

168 The following subsections describe the ozone data sources in detail.

169

170 *3.1 SBUV*

171 SBUV is a nadir-viewing instrument that measures incoming solar irradiance and

172 backscattered solar radiance reflected from the Earth's atmosphere directly beneath the

173 satellite. Measurements are made in 12 narrow wavelength bands between 250 and 340 nm

174 using a double monochromator with a 1.1 nm (full-width half maximum) triangular slit

175 function (Fleig et al., 1990). The ratio of radiance to irradiance at various wavelengths is

176 used to infer the amount of ozone in broad layers of the atmosphere in the column below the

177 instrument. The total column ozone is the sum of the profile layer values. Satellite

178 measurements of stratospheric ozone from the SBUV instrument began in November 1978

179 with the launch of the Nimbus-7 spacecraft. A series of SBUV/2s began operations with the

180 launch of NOAA-9 in January 1985 and continued with the sequential launches of NOAA-

181   11, -14, -16, -17, -18, and -19, with NOAA-19 still operational at the time of publication.

182   The MERRA-2 processing ingests data from Nimbus-7 and NOAAs 11-17, following the

183   time line in Table 1, before switching to data from EOS Aura. We include measurements

184   taken at solar zenith angles less than 84 degrees. Measurements are available nearly every

185   day, with the exception of one month of missing data in March 1991 from NOAA-11

186   SBUV/2. Hereafter we refer to SBUV and SBUV/2 simply as 'SBUV'.

187

188   SBUV measurements are made every 32 seconds, giving approximately 100 measurements

189   per orbit (about 1 every 1.85 degrees in latitude; each orbit is separated by ~ 26 degrees

190   longitude), or over 1400 measurements on a given day. The instrument field of view traces a

191   188 km wide swath (200 km for Nimbus-7 SBUV) along the orbital track. All the data have

192   been processed with the Version 8.6 retrieval algorithm (V8.6; Bhartia et al. 2013), an

193   update to the Version 8 data assimilated in MERRA. In the V8.6 algorithm, the ozone

194   cross-sections are taken from (Daumont et al. 1992), which are superior in resolution,

195   temperature dependence, and quality to the Bass and Paur (1985) cross-sections used in

196   prior retrievals. A more accurate cloud height climatology has been developed using the UV

197   rotational Raman filling technique (Vasilkov et al. 2004) from the OMI onboard EOS-Aura.

198   The data set contains the climatological heights of the "optical centroid pressure," which

199   reflect how deep a UV photon, on average, will penetrate into the cloud. With this more

200   accurate cloud height climatology, the errors produced by extrapolating ozone amounts

201   under a cloud are minimized.

202

9

203 Instrument noise for the majority of measurements is less than 0.5% (Deland et al., 2012).

204 Bhartia et al. (2012) describes the V8.6 algorithm and related uncertainties in detail. The

205 largest source of error in the SBUV profile retrievals is the smoothing error (Bhartia et al,

206 2012; Kramarova et al., 2012). The smoothing error describes the component of vertical

207 ozone variability which the observation system cannot measure. Between 16 and 1 hPa, the

208 smoothing errors for SBUV monthly zonal mean retrievals are of the order of 1%,

209 increasing to 15–20% in the troposphere. The smoothing errors for total ozone retrievals are

210 mostly less than 0.5%. In MERRA-2 the layer-dependent SBUV partial column error

211 specification arises from tuning performed at the National Centers for Environmental

212 Prediction and follows other GMAO operational analyses that use SBUV: the errors are

213 assumed to be 1.4 Dobson units (DU, ~15%) between 1000 hPa and 631 hPa, 1.9 DU (11 %

214 - 6 %) in the lower stratospheric layers, 100 hPa – 63 hPa and 63 hPa – 40 hPa and

215 decreasing to 1 DU (25 % to over 100 %) in the upper stratospheric and mesospheric layers.

216 Note that with these error specifications SBUV data have little impact in the upper

217 stratosphere and mesosphere. The total ozone error is set to 6 DU (about 2 % of the global

218 total ozone) in agreement with ground-based and ozonesonde comparisons done by Labow

219 et al. (2013).

220

221 The SBUV series of instruments were calibrated using a combination of "hard calibration"

222 from pre-launch and on-orbit instrument monitoring, and a variety of "soft calibration"

223 techniques as described in Deland et al. (2012). In addition, the instruments are cross-

224 calibrated to each other at the radiance level within the V8.6 retrieval algorithm.

225 Nevertheless differences between the instruments exist and the data quality varies from one

226 instrument to the next. In particular, measurements from the NOAA-14 SBUV/2 instrument

227 show enhanced uncertainty relative to the other instruments in validation studies

228 (Kramarova et al. 2013b).

229

230 *3.2 OMI*

231 The Ozone Monitoring Instrument (Levelt et al. 2006) operates on NASA's EOS Aura

232 satellite, which was launched on 15 July 2004 into a sun-synchronous orbit with a 1:45 PM

233 equatorial crossing time on the ascending node. The instrument is equipped with a 60-pixel

234 cross-track sensor array measuring backscattered solar radiation in the 270-550 nm

235 wavelength range with a spectral resolution of ~0.5 nm at nadir. The treatment of OMI

236 observations in MERRA-2 closely follows that of Wargan et al. (2015) and we will only

237 summarize it briefly here. The reanalysis assimilates total column ozone data using the

238 version-8.5 retrieval algorithm extensively evaluated by McPeters et al. (2008). Of the 60

239 available pixels (rows), only ozone columns from rows 3-24 are used. This is motivated by

240 the lower quality of the data from rows affected by a mechanical issue known as the row

241 anomaly from 2008 onward as well as large pixel sizes from rows 1 and 2. With this

242 selection the swath width is about 1,100 km. The total column ozone from OMI is

243 assimilated using efficiency factors provided with the data in order to account for the lower

244 sensitivity of OMI measurements in the lower stratosphere, specifically in clouded scenes.

245 This methodology is described in detail in Wargan et al. (2015).

246

247 The following changes to the OMI treatment were applied in MERRA-2 relative to the

248 approach of Wargan et al. (2015) following recommendations of the OMI science team

249   (P.K. Bhartia, personal communication): (1) Only the pixels with the total ozone quality flag

250   set to 0 are used and (2) The efficiency factors are set to 1.0 above 125 hPa. MERRA-2

251   assumes a constant observation error of 5 DU everywhere. This number is within 2 % of the

252   globally averaged total ozone and it is consistent with the level of agreement between OMI

253   and ground-based measurements reported by McPeters et al. (2008). McPeters et al. (2015)

254   note that OMI exhibits a latitude-dependent low bias with respect to SBUV (from 0.5 %

255   between 60°S and 15°S to 2 % near 60°N). At present, the source of this bias is not

256   understood and no attempt was made in MERRA-2 to remove it. It leads to an increase of

257   the difference between SBUV data and MERRA-2 total ozone from 1 % to 2 % between the

258   tropics and northern high latitudes, as discussed below in Section 6.1.

259

260   *3.3 Aura MLS*

261   The Microwave Limb Sounder on the EOS Aura satellite (Waters et al. 2006) measures

262   profiles of atmospheric thermal radiation in a broad spectrum of microwave bands allowing

263   high quality retrievals of temperature and concentrations of over a dozen chemical species in

264   the stratosphere, including ozone. It provides about 3,500 retrieved stratospheric ozone

265   profiles daily covering the 82°S-82°N latitude range in both day and night. The vertical

266   resolution, determined from the full width at half maximum of the averaging kernels of the

267   MLS ozone data, ranges from 2.5 km in the middle stratosphere to 6 km in the mesosphere

268   (Froidevaux et al. 2008). Note that this is not the same as the spacing of the nominal levels

269   on which the data are provided (see below). MERRA-2 uses version 2.2 ozone retrievals

270   (Froidevaux et al. 2008) between October 2004 and May 2015 and version 4.2 (Livesey et

271   al. 2015) afterwards. We note that MERRA-2 was already in production when version 4.2

272 data were released so the use of an older version was necessary in the preceding period. The

273 decision to use version 2.2 rather than 3.3 as in Wargan et al. (2015) was motivated by the

274 occurrence of unphysical vertical oscillations in the tropical profiles seen in version 3.3 that

275 were present but to a much lesser degree in version 2.2. In both versions ozone is retrieved

276 from the 230–250 GHz spectral region but the nominal vertical resolutions and ranges differ

277 between the different versions. The recommended vertical range for version 2.2 is 215 hPa –

278 0.02 hPa. However, following Wargan et al. (2015) we applied mid-layer averaging to the

279 profiles prior to assimilation, consistent with the fact that the GEOS-5 ozone represents

280 layer-averaged concentrations so that the lowest assimilated layer is centered on 177.8 hPa

281 and the highest at 0.05 hPa, 21 levels total. In version 4.2 the number of MLS levels in the

282 stratosphere increased by almost a factor of two, compared to version 2.2. Mid-layer

283 averaging is not applied to these high-resolution data because the version 4.2 profiles are

284 sufficiently smooth and we verified that averaging has little effect on the input data.

285 Assimilating version 4.2 ozone at the nominal levels allows us to extend the vertical range

286 of the input MLS profiles from 261 hPa (215 hPa since May 2016) to 0.02 hPa on 36 levels.

287 However, preliminary comparisons indicate a high bias in the 261 hPa level compared to

288 ozonesondes. Starting 1 May 2016 that level was turned off in MERRA-2 so that the lowest

289 assimilated MLS level is 215 hPa from that date onward. MLS provides an almost unbroken

290 record of observations from 2004 to present. The longest data gap occurs between 27 March

291 and 18 April 2011 due to technical issues with the instrument.

292 The assumed observation errors are calculated from the square root of the sums of squares

293 of precision (provided with the data) and 0.5 times the accuracy estimations from the MLS

294 data quality documents (Livesey et al. 2007 for version 2.2 and Livesey et al. 2015 for

295 version 4.2). The factor of 0.5 is applied because the accuracy estimates are quoted as a

296 notional "2-sigma" term. The data selection is done before assimilation and follows the

297 guidelines given in the MLS data quality documents.

298 This study evaluates the MERRA-2 ozone product between the years 1980 and 2013 when

299 sufficient number of independent data (in particular from ozonesondes) are readily available

300 to us.

301 *3.4 Spatial coverage of the ozone data*

302 Figure 1(a) shows the time series of the global monthly total ozone observation counts from

303 data sources assimilated in MERRA-2. Each SBUV instrument except the sounder on

304 Nimbus-7 provided between 30,000 and 40, 000 total ozone observations per month. The

305 Nimbus-7 SBUV had two periods of slightly lower data counts: 1980-1983 and 1987-1990.

306 From 1980-1983 SBUV operated on a 3-day on 1-day off cycle to reduce power load on the

307 satellite, but as other instruments failed power constraints were reduced allowing SBUV to

308 operate continuously after mid-1983. Nimbus-7 SBUV data after February 1987 are affected

309 by chopper wheel synchronization errors. This caused an increase in measurement noise and

310 a reduction in the number of usable profiles from SBUV over the remainder of the data

311 record (Gleason and McPeters 1995). The monthly number of OMI observations (2004

312 onward) is about an order of magnitude greater.

313 Panel (b) of Figure 1 plots the monthly latitudinal coverage of each total ozone data type

314 used. Besides a limited extent of the data during winter, the coverage of the SBUV

315 instruments on NOAA-11 and NOAA-14 suffers from the effects of orbital drifts of these

316 spacecraft. In particular, the equatorial crossing time of NOAA-11 changed from about 2PM

14

317 in 1989 to 5PM in 1994, severely limiting SBUV coverage. By mid-1994 ozone

318 observations were unavailable south of 30°S. Similarly, there was a loss of high-latitude

319 coverage in 2001 caused by the drift of the NOAA-14 satellite. The MERRA-2 ozone

320 product in the high latitudes in these years should be not be used. The OMI instrument

321 provides steady coverage of the sunlit atmosphere from 2004 onward with the exception of

322 an outage between 29 May and 13 June 2016 (beyond the time period considered here).

323 The coverage of partial column ozone observations from SBUV (1980-2004) is the same as

324 for the total ozone (in the SBUV period total ozone data are simply the sum of the SBUV

325 layer values). The geographical coverage of MLS is 82°S-82°N throughout the period of the

326 reanalysis.

327 **4. Changes from MERRA relevant to ozone**

328 The main difference in the treatment of ozone between MERRA and MERRA-2 is in the

329 observing systems that the two reanalyses use. The ozone observations assimilated in

330 MERRA were version 8 retrieved partial and total columns from SBUV. The partial

331 columns were mapped from 21 to 12 layers and the total ozone column data were obtained

332 by summing the layer values in each profile. By contrast, prior to October 2004 MERRA-2

333 uses the most recent version 8.6 of SBUV data, which is expected to provide a more realistic

334 continuity across different satellite platforms as discussed in Section 3.1. These data are

335 assimilated on the native 21 layers, although one should keep in mind that significant

336 smoothing errors exist in the vertical (Kramarova et al. 2013). After 1 October 2004,

337 MERRA-2 assimilates OMI and MLS data, which were not included in MERRA.

338

339   In data assimilation, the information from observations is applied to the background fields

340   provided by the GCM forecast using background error covariances which control how the

341   data information is propagated and optimally combined with the background in the

342   horizontal and vertical directions (Cohn 1997; Lahoz et al. 2007). Following Wargan et al.

343   (2015), the MERRA-2 data assimilation algorithm assumes that the background error

344   standard deviation for ozone at any given grid point and time is proportional to the

345   background concentration. Test experiments demonstrated that this leads to a more realistic

346   representation of shallow vertical structures in the ozone profiles in the UTLS than when

347   static errors, such as those employed for MERRA, are used.

348

349   The models in both reanalyses use the same set of monthly 2-dimensional ozone production

350   rates and loss frequencies derived from a 2-dimensional chemistry model as described by

351   Stajner et al. (2008) and Wargan et al. (2015). As argued in Wargan et al. (2015), such a

352   simplified chemistry scheme is sufficient for a data-driven assimilated product in the upper

353   troposphere and lower stratosphere where the bulk of ozone is located and where chemical

354   time scales range from weeks to months – very long compared to the frequency of data

355   insertion (daily). However, in the absence of day and night observations, prior to October

356   2004, the diurnal cycle in the upper stratospheric ozone is not represented in MERRA-2.

357

358   **5. Independent ozone data**

359

360   This section describes the data used for comparisons with MERRA and MERRA-2. None of

361   the observations listed below are assimilated in either reanalysis.

362

5.1 TOMS

364   The Nimbus-7 Total Ozone Mapping Spectrometer (TOMS) instrument is a downward-viewing

365   spectrometer, which measures Earth-backscattered UV radiances (Herman et al. 1991).

366   Radiation in six discrete 1-nm wavelength bands (312.5, 317.5, 331.2, 339.8, 360, and 380 nm)

367   is measured at 35 cross-track scan positions. Measurements from successive orbits overlap

368   resulting in daily global coverage. The spatial resolution is 50 km by 50 km for the nadir view,

369   and about 50 km by 200 km at the extreme cross-track scan positions. The TOMS instrument

370   also measures the solar irradiance for each wavelength every day using a diffuser plate to reflect

371   sunlight into the instrument. These solar irradiance measurements provide radiance

372   normalization and remove some instrumental dependence. Total column ozone (TCO) data are

373   derived from the Version 8 TOMS algorithm (Bhartia and Wellemeyer 2002; Wellemeyer et al.

374   2004), which was released in 2004. The algorithm uses only two wavelengths (317.5 and 331.2

375   nm) to derive TCO data; the other four wavelengths are used for error correction and

376   identification of aerosols and clouds. Uncertainty estimates for Nimbus-7 TOMS total ozone are

377   +3 % for absolute error, ±2 % (1 sigma) for random error, and +1.5% for the drift over 14 years

378   with values somewhat larger at higher latitudes (McPeters et al. 1996).

379

5.2 Ozonesondes

381   The ozonesonde observations are as in Wargan et al. (2015): the data are from the Network

382   for the Detection for Atmospheric Composition Change, the Southern Hemisphere

383   Additional Ozonesondes (SHADOZ: Thompson et al. 2003a), as well as field campaigns.

384   See Wargan et al. (2015) for a discussion of the ozonesonde locations, precision and

385 accuracy and representativeness issues inherent in comparisons of point measurements with

386 gridded analysis data.

387

388 In addition we compare MERRA-2 against the TCO data derived from the ozonesonde

389 observations at the South Pole. Since the vertical range of balloon-borne observations

390 typically does not extend above 10 hPa, the upper stratospheric contribution to the TCO

391 provided with the ozonesonde data files is calculated assuming the constant mixing ratio

392 above 7 hPa or at the altitude of balloon burst if that occurred below 7 hPa. The South Pole

393 ozonesonde data was downloaded from the Earth System Research Laboratory website

394 (http://www.esrl.noaa.gov/gmd/ozwv/ozsondes/spo.html).

395

396 *5.3 UARS MLS*

397 The Microwave Limb Sounder instrument on the Upper Atmosphere Research Satellite

398 (UARS MLS: Barath et al. 1993) operated between September 1991 and July 1999 but the

399 observation frequency declined significantly after 1993. In this work we use UARS MLS

400 observations between 1991 and 1996. The spatial coverage alternated from month to month

401 between 34°S-80°N and 80°S-34°N due to the spacecraft's monthly yaw maneuvers. We use

402 Version 5 ozone profiles retrieved from UARS MLS measurements of the atmospheric limb

403 emissions at 205 GHz described and validated by Livesey et al. (2003). All profiles were

404 interpolated to a set of constant pressure levels, 16 per decade pressure. There are 29 UARS

405 MLS levels between 60 hPa and 1 hPa (the vertical range used in our comparisons). We

406 have applied a 2-point vertical boxcar smoother to all UARS profiles in order to reduce

407 unphysical oscillations resulting from the reported vertical resolution being higher than that

408 resulting from the actual information content.

409

410 *5.4 MIPAS*

411 The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) instrument was

412 flown on the European Space Agency's (ESA) Envisat satellite, launched on 1 March 2002.

413 It was a middle-infrared Fourier transform spectrometer measuring limb emissions in the

414 4.15 μm - 14.6 μm spectral range (Fischer et al. 2008). The instrument operated in its

415 standard observation mode (nominal spectral resolution of 0.025 cm^{-1}) between July 2002

416 and March 2004. The operations were halted until the end of 2004 due to technical problems

417 and resumed in January 2005 at a reduced spectral resolution of 0.0625 cm^{-1} and different

418 selection of spectral microwindows, but with denser spatial sampling (Cortesi et al. 2007;

419 Raspollini et al. 2013). The data record extends to early April 2012 when communication

420 with Envisat was permanently lost. We use the ozone product from version 6 of the ESA

421 retrieval algorithm described by Raspollini et al. (2013). The latter study concludes that the

422 quality of the retrieved species is generally better for the 2005-2012 period. In particular the

423 vertical resolution of ozone retrievals in the first period (2002-2004) ranges from 3.5 km to

424 5.5 km between 20 km and 50 km, but in the latter period ranges from 2.5 km to 4 km

425 (Raspollini et al. 2013, their Figure 9). MIPAS profiles are interpolated to a 26-level vertical

426 grid with 12 levels between 60 hPa and 1 hPa (note that this is different than for UARS

427 MLS).

428

429 *5.5 SAGE II*

430 The Stratospheric Aerosol and Gas Experiment II (SAGE II) is a solar occultation

431 instrument flown on the Earth Radiation Budget Satellite (ERBS) between October 1984

432 and August 2005 measuring the atmospheric transmission of the solar radiation in 7

433 channels nominally located between 1020 nm and 386 nm. The number of profiles is

434 reduced by a factor of 2 after the year 2000. The occultation methodology provides high

435 vertical resolution and high precision retrievals of stratospheric ozone (along with water

436 vapor, NO_2 and aerosols) but limits the data coverage to 30 profile measurements per day,

437 15 at sunrise and 15 at sunset for each orbit. The latitudinal coverage varies from month to

438 month with the extent rarely exceeding about 60°S to 60°N. The retrieved SAGE II ozone

439 profiles have been used extensively in stratospheric ozone studies, as summarized in WMO

440 (2014).

441

442 We use Version 7 of the retrieval algorithm described and evaluated by Damadeo et al.

443 (2013). The profiles are interpolated onto a 61-level vertical grid between the surface and

444 0.2 hPa. Comparisons shown in this study use the SAGE II data at pressures lower than

445 about 200 hPa.

446

447 **6. Results of the Comparisons**

448

449 This Section describes comparisons of MERRA-2 ozone fields against independent data

450 from TOMS (total ozone); SAGE II, UARS MLS, MIPAS (stratospheric profiles) and

451 ozonesondes (UTLS and integrated profiles). The comparisons are done as follows: for each

452 independent observation, the closest MERRA-2 (and MERRA, if applicable) geolocation is

453   found. Then, for profile data, the reanalysis mixing ratio profile is interpolated linearly in

454   log-pressure to the verifying observation levels. Unless stated otherwise, all the statistics

455   considered here are calculated from pairs of data collocated in this way. We compute spatial

456   and temporal means to identify biases and standard deviations and correlations to evaluate

457   variability.

458

459   *6.1. Total ozone – comparisons with TOMS and South Pole ozonesondes*

460

461   Comparisons of MERRA and MERRA-2 TCO fields against TOMS data (January 1980 –

462   May 1993) are done as follows. First, TOMS level 2 data are mapped onto a 1°×1° grid by

463   averaging all measurements taken within each grid-box. The gridded data are then matched

464   with the MERRA and MERRA-2 TCO in space and time (within 1 hour of the average time

465   of measurements in the grid box) and monthly statistics are computed.

466

467   Figure 2 shows the time series of monthly TCO statistics broken down into 5 latitude bands:

468   90°S-60°S, 60°S-30°S, 30°S-30°N, 30°N-60°N and 60°N-90°N. The results are summarized

469   in Table 2. Monthly mean percent differences (the left column in Figure 2, panels a, c, e, g

470   and i) show that with the exception of the southern high latitudes (panel a) after 1990 the

471   agreement between both reanalyses and TOMS is within 5 %, with some seasonal

472   dependence outside of the tropics. Overall MERRA-2 is about 1 % lower than MERRA.

473   This is consistent with the fact that the total ozone derived from the version 8.6 of SBUV

474   algorithm is slightly lower than in version 8 used in MERRA (McPeters et al. 2013). The

475   MERRA-2 minus TOMS total ozone difference averaged over the period of comparison is

476   shown in Table 2. The reanalysis is slightly higher than TOMS in high latitudes (up to 1.45

477   % in the northern high latitudes) and lower by up to 1.83 % in the middle latitudes and the

478   tropics. While a detailed discussion of MERRA-2 assimilation internal statistics is beyond

479   the intended scope of this study we note that the total ozone tendencies due to dynamics in

480   the GEOS-5 DAS are generally negative in the tropics and positive in high latitudes,

481   suggesting that the meridional overturning (Brewer-Dobson) circulation from the

482   assimilated meteorological fields is too rapid. This is in part countered by the tendencies

483   arising from ozone data insertion, whose sign is the opposite of the dynamical tendencies on

484   the average. The end result is a compromise between the tendencies due to dynamics and

485   data assimilation, consistent with the pattern of positive (negative) bias in high (low-middle)

486   latitudes seen in Table 2.

487

488   Starting in 1991 the agreement between the reanalyses and TOMS is degraded in the 90°S-

489   60°S latitude band (Figure 1a). This is an expected result of the diminished coverage of the

490   NOAA-11 SBUV in the southern high latitudes.

491

492   The right column of Figure 2 (panels b, d, f, h and j) shows the time series of monthly

493   standard deviations of the reanalysis minus TOMS differences. The same statistics for

494   MERRA-2 averaged over the period of comparison are summarized in Table 2. The

495   difference standard deviations are almost identical for both reanalyses, ranging from about

496   2.5 % in the tropics (Figure 2f) to 8 % in high latitudes during spring (Figure 2b and j). The

497   multiyear average for MERRA-2 is less than 5.5 % almost everywhere (Table 2),

498   comparable with the assumed observation error for total ozone data and indicating excellent

499   agreement of the reanalysis with TOMS. Here again, the agreement deteriorates in the

500   southern high latitudes after 1991 (Figure 2b). We note three spikes in the reanalysis minus

501   TOMS standard deviation and mean time series in the tropics (Figure 2e and f), one in 1982

502   and two in 1991. Two of these spikes, denoted by the dashed lines in the figure, correspond

503   to the eruptions of El Chichon and Mt. Pinatubo. High stratospheric aerosol loading

504   following large volcanic eruptions degrades the quality of ozone retrievals from UV

505   measurements (Bhartia et al. 2013). Shorter wavelength measurements (SBUV profile) are

506   more affected than longer wavelength measurements (TOMS TCO) when aerosol levels are

507   elevated above the ozone peak, leading to a different response between SBUV and TOMS

508   (Torres et al. 1995; Torres and Bhartia 1995). Averaging over multiple scan angles further

509   reduces the aerosol signal in TOMS (Torres et al. 1995). The third spike, denoted by the

510   solid line, results from a full month of missing SBUV data (see Section 3.1).

511

512   When considering total ozone column, it is instructive to compare MERRA-2 with the

513   SBUV-derived Merged Ozone Data set (MOD: Frith et al. 2014) based on version 8.6 of

514   SBUV retrievals, the same as MERRA-2 uses. Figure 3 plots the relative difference between

515   MERRA-2 total ozone and MOD data as a function of latitude averaged over two time

516   periods: 1980-2003 (SBUV period) and 2005-2014 (EOS Aura period). The comparison

517   serves as a consistency (MERRA-2 assimilates SBUV data in the former period) and

518   continuity check between the two periods. In the SBUV period MERRA-2 is slightly lower

519   than MOD (by up to 1 %) and the difference is approximately symmetric with respect to the

520   equator. By contrast, the low bias in the EOS Aura period exhibits a clear latitude

521   dependence: it is in agreement with the MERRA-2 − MOD difference in the southern

522 hemisphere but in the northern hemisphere the absolute value of the bias increases linearly

523 with latitude reaching ~2 % at 80°N. This is consistent with the latitude-dependent low bias

524 in the OMI data (Section 3.2).

525

526 Figure 4 shows comparisons of total ozone from MERRA and MERRA-2 at the South Pole

527 with values derived from ozonesonde measurements. All three datasets show pronounced

528 annual variations (Figure 4a) with sharp minima reaching down to 100 DU during ozone

529 hole conditions (September-October) followed by maxima in austral summers and periods

530 of decline leading up to the next year's ozone hole. Panel (b) of Figure 4 shows the

531 reanalysis minus ozonesonde differences relative to individual total ozone values derived

532 from the ozonesonde profiles. Aside from large positive excursions during the austral

533 winter-spring transitions (with limited UV data coverage) occasionally reaching over 100 %,

534 the differences are within 20 % for both reanalyses. Statistics calculated from these

535 comparisons show that during both the SBUV and Aura periods MERRA-2 is lower than the

536 sondes by 3 %. The standard deviation of MERRA-2 minus ozonesonde differences is 12.47

537 % in the SBUV period and only 5 % in the Aura period. The corresponding standard

538 deviation numbers for MERRA are 13.3 % and 15.9 %, respectively, indicating a much

539 better performance of MERRA-2, especially in the period when MLS data are assimilated.

540 We note the realism of the interannual variability in MERRA-2. In particular, the record

541 high springtime ozone associated with the major sudden stratospheric warming in

542 September 2002 is well represented by both MERRA and MERRA-2 (Figure 4a).

543

544   Overall, the total ozone product in MERRA-2 compares well with independent data. While

545   small systematic season-dependent biases exist, the annual cycle, latitudinal structure and

546   longer-term variability are realistic and the agreement with the independent data is well

547   within the assumed observation errors.

548

549   *6.2 Comparisons with SAGE II*

550

551   Comparisons of MERRA and MERRA-2 against SAGE II observations are shown in

552   Figures 5-8 and Tables 3 and 4 for the period 1984-2005, though the number of SAGE

553   profiles is reduced by a factor of two after 2000. Note that most of the SAGE II record falls

554   into the SBUV period for MERRA-2.

555

556   Figure 5 shows the time series of annual relative mean ozone differences integrated between

557   208 hPa and 0.2 hPa (hereafter, stratospheric column) in three broad latitude bands: south of

558   30°S, between 30°S and 30°N, and north of 30°N. MERRA and MERRA-2 minus SAGE II

559   differences are shown. Note that the latitudinal extent of SAGE II observations is about

560   60°S-60°N so that the comparisons do not cover the polar regions. The years 1992 and 1993

561   are omitted because of high aerosol loading affecting ozone retrievals in the lower

562   stratosphere in years following the eruption of Mt. Pinatubo (1991). The differences are

563   within 2 % for both reanalyses and exhibit an upward drift over the period of comparison in

564   the southern hemisphere and in the tropics. Note the large difference between MERRA and

565   SAGE II south of 30°S in 1994 due to limited SBUV coverage (see Figure 1). This effect is

566   much reduced in MERRA-2. For most of the comparison period MERRA-2 is slightly

567   higher than SAGE II in all latitude bands. Wang et al. (2002) report that version 6.1 of

568   SAGE II ozone data exhibit a low bias with respect to ozonesondes, which is most

569   pronounced in the tropics below 15 km (~125 hPa) but also present at the middle latitudes

570   (see their Figures 9 and 10). While Version 7 of the SAGE II data used here is much

571   improved in the middle to upper stratosphere compared to previous versions (Damadeo et al.

572   2013), the differences in the lower stratosphere are smaller and it can be expected that some

573   of the bias between the reanalyses and SAGE II stratospheric column is due to an

574   underestimate of the lower stratospheric ozone in SAGE II retrievals.

575

576   Figure 6 (a-c) plots the time series of annually averaged zonal mean ozone from SAGE II

577   and both reanalyses south of 30°S at three SAGE II levels: 42.7 hPa (lower stratosphere),

578   10.1 hPa (middle stratosphere) and 4.3 hPa (upper stratosphere). The statistics averaged

579   over the period of comparison (1984-2005) are given in Table 3 for MERRA-2 only. At

580   42.7 hPa the MERRA-2 – SAGE II differences are small (2 %) and both reanalyses follow

581   the interannual variability seen in SAGE II observations. The same is true at 10.1 hPa with

582   the exception of 1994 when MERRA is ~0.4 parts per million by volume (ppmv),

583   equivalently 6 %, higher than SAGE II. The bias in MERRA-2 is about half of that value.

584   As seen in Figure 1 the NOAA-11 SBUV coverage was very limited in that year, especially

585   in the southern hemisphere. At 4.3 hPa both reanalyses are very close to SAGE II between

586   1985 and 1991. Between 1992 and 2004 MERRA is slightly closer to SAGE II

587   measurements than MERRA-2.

588

589   Panels (d)-(f) of Figure 6 show the standard deviations of the SAGE II minus reanalysis

590   differences at the same levels. Table 3 lists the MERRA-2 results averaged between 1984

591   and 2005. The results for both reanalyses are close to each other with the exception of 2005

592   (the first full year of the Aura data assimilation) at 42.7 hPa where the agreement of

593   MERRA-2 with SAGE II improves by almost a factor of 2. The MERRA-2 – SAGE II

594   difference standard deviations range from 0.31 ppmv (4.6 %) at 4.3 hPa to 0.35 ppmv (11

595   %) at 42.7 hPa.

596

597   We will now describe briefly the results for the northern middle latitudes (north of 30N°, not

598   shown). They are qualitatively similar to those for the southern hemisphere. Of note is the

599   fact that after 1995 MERRA-2 has a larger positive bias with respect to SAGE II at 4.2 hPa

600   but captures more of the variability within that year. A large drop in the difference standard

601   deviation at 42.7 hPa in 2005 for MERRA-2 is seen in both hemispheres, though the number

602   of SAGE profiles is limited in 2005. The results for the northern middle latitudes are

603   summarized in Table 4.

604

605   The bottom row (panels g-i) of Figure 6 shows the annual SAGE II data counts for each

606   level and hemisphere. In the middle and upper stratosphere there are close to 4,000

607   observations per year until the year 2000 and about 2,000 afterwards. The data counts in the

608   lower stratosphere (at 42.7 hPa) are slightly above 2,000 before the year 2000 except in the

609   years following the Mt. Pinatubo eruption when the numbers drop to less than 1,000.

610

611   Sunrise ozone mixing ratios from SAGE II tend to be lower by 8–10 % relative to sunset

612   data above 35 km owing to diurnal variations in upper stratospheric ozone and a possible

613   contribution from algorithmic biases (Kyrölä et al. 2013; Damadeo et al. 2014; Sakazaki et

614   al. 2015). Tables 3 and 4 include statistics calculated from MERRA-2 and SAGE II for

615   sunset-only measurements (in parentheses). While the difference standard deviations

616   calculated for the sunset-only data are very close to the all-data values the biases are much

617   lower, especially in the upper stratosphere. For example, at 4.2 hPa in the northern

618   hemisphere the MERRA-2 bias drops from 1.1 % to 0.3 % if only sunset observations are

619   used.

620

621   In Figure 7 we show the mean and standard deviation of differences between the reanalyses

622   and SAGE II profiles relative to the SAGE II mean from all available data between January

623   and August 2003 (the SBUV period) and for the same months in 2005 (the Aura period).

624   The results are given for three latitude bands: south of 30°S, between 30°S and 30°N, and

625   north of 30°N. Both reanalyses exhibit a low bias of as much as to 5 % between 5 hPa and 1

626   hPa, but the bias is reduced in MERRA-2 in 2005 compared to 2003. In 2003 MERRA-2

627   has an alternating pattern of low and high bias with respect to SAGE II in all latitude bands.

628   By contrast, in 2005 there is almost no bias between MERRA-2 and SAGE II in the middle

629   stratosphere (10-50 hPa) in the extratropics and only a small (up to 3 %) negative bias

630   around 10 hPa in the tropics. The difference standard deviation ranges from 10 % - 30 %

631   between 100 hPa and 60 hPa (30 % in the tropics near 100 hPa) to less than 10 % in the

632   middle stratosphere. While in 2003 (the SBUV period) the standard deviation profiles are

633   almost identical between MERRA and MERRA-2, the latter better captures the SAGE II

28

634 variability in 2005 when stratospheric ozone is constrained by higher vertical resolution

635 MLS observations. SAGE II – MERRA-2 correlation profiles (not shown) also improve in

636 2005 compared to 2003. South of 30°S the correlation increases from ~0.85 to ~0.95 at

637 pressures higher than 20 hPa with the largest improvement at 40 hPa where the correlations

638 are 0.77 and 0.94 for 2003 and 2005, respectively. In the middle stratosphere the

639 correlations are similar for both years (~0.95) with larger differences at pressures lower than

640 2 hPa. Similar improvement is seen north of 30°N. In the tropics, the correlations also

641 increase by about 0.1 throughout the stratosphere with the 2005 values between 0.8 and

642 0.95, except in the upper stratosphere (the correlation is 0.5 at 1 hPa in 2005). These results

643 confirm a good representation of vertical structures in the MERRA-2 stratospheric ozone

644 and a pronounced further improvement in the Aura era.

645

646 *6.3 Comparisons with UARS MLS*

647

648 Because of the limited time span of UARS MLS measurements (days when data are

649 available become sparse after 1993, we will only analyze the relevant statistics aggregated

650 over the 1991 – 1996 period and refrain from discussing time-dependent statistics. Note that

651 the years of UARS MLS operations fall entirely within the SBUV period of MERRA-2.

652

653 The results are shown in Figure 8 separately for 4 seasons: December-February (DJF),

654 March-May (MAM), June-August (JJA), and September-November (SON). The top row

655 shows the zonal mean MERRA-2 minus UARS MLS difference as a percent of the average

656 UARS MLS ozone at each level between 60 hPa and 1 hPa for each of the seasons. Overall,

657   MERRA-2 is lower than UARS MLS. In most of the stratosphere the pattern is similar in

658   each season and consists of a low bias layer (4 % - 12 %) between 3 hPa and 1 hPa, and

659   between 10 hPa and 30 hPa, smaller high bias between 50 hPa and 30 hPa (except in high

660   latitudes where the difference is negative) and again low bias below 50 hPa. The latter is

661   large (over 20 %) but one must remember that the mean ozone concentrations are small in

662   the lower stratosphere, contributing to a large relative difference. The MERRA-2 minus

663   UARS MLS difference between 100 hPa (not shown) and 50 hPa does not exceed ~0.2

664   ppmv. Also seen in Figure 8 are regions of positive differences in the high latitudes between

665   10 hPa and 2 hPa in the northern hemisphere during DJF, in the southern hemisphere during

666   MAM and JJA, and in both hemispheres during SON. The largest differences of up to 25 %

667   are seen in JJA at about 10 hPa in the southernmost latitudes.

668

669   The middle row of Figure 8 (panels e-h) show the standard deviation of the MERRA-2 –

670   UARS MLS difference expressed as a percent of the average UARS MLS ozone. Also

671   plotted is the zonal standard deviation of UARS MLS ozone calculated for the same periods

672   as a measure of the ozone fields' variability (white contours). Broadly speaking, the

673   difference standard deviations and the variability contours exhibit very similar patterns with

674   the former being always less than the latter, indicative of a realistic representation of the

675   stratospheric ozone variability in MERRA-2. In most of the stratosphere the difference

676   standard deviation is much less than 10 %. Higher values of up to 25 % are seen only in

677   regions of high variability such as the high latitudes in southern hemisphere winter and

678   spring, where SBUV measurements are limited, and in the lower stratosphere in the tropics,

679   where ozone concentrations are low.

680

681   Correlations between MERRA-2 and UARS MLS are shown in panels (i)-(l) of Figure 8.

682   The zonal patterns follow that of the ozone variability with the highest correlations of 0.7

683   and higher in the extratropics (somewhat reduced during summer). The lowest, even

684   negative correlations are found in the tropics, where they exhibit a vertical pattern

685   resembling the QBO signature in ozone. Coy et al. (2016) showed that the ozone response to

686   the QBO in MERRA-2 does not exhibit a realistic vertical structure during the SBUV

687   period. This is a result of a limited vertical resolution (large smoothing errors) in SBUV

688   profile data as explained by Kramarova et al. (2012).

689

690   We want to emphasize that despite the differences between MERRA-2 and UARS MLS, the

691   two data sets agree quite well in terms of the structure of the ozone fields. As an example,

692   Figure 9 shows the ozone field at 10 hPa on 21 December 1993. A breaking planetary wave

693   pulled a thin tongue of ozone-rich subtropical air from over the coast of East Asia into the

694   high latitudes, transporting it along the edge of the polar vortex almost over the North Pole

695   and wrapping it around the Aleutian anticyclone. This is a fairly common occurrence in the

696   winter polar stratosphere. In this case MERRA-2 shows that the polar vortex was displaced

697   towards Europe. The 10 hPa temperature over the pole increased rapidly by ~40°K during

698   the final days of December (not shown), indicative of a sudden stratospheric warming.

699   Overlaid on the MERRA-2 ozone field in Figure 9 are UARS MLS observations taken

700   within 4.5 hours of 6Z, the time of the analysis. There is considerable agreement between

701   the data sets: MLS samples the low mixing ratio inside the polar vortex, the tongue of high

702   ozone from Kamchatka into the Arctic Ocean, the region of high values over Canada, and

703   the high ozone filament between 30°N and 40°N from the North Atlantic to north of East

704   Siberia. We note that the NOAA-11 SBUV coverage extended only up to ~51°N at that time

705   (marked by black circles in Figure 9(a)) so that all the information about the ozone

706   distribution over the polar region in MERRA-2 is generated by the general circulation

707   model, in particular via advection of air masses from lower latitudes where observations are

708   available. Figure 9(b) shows a scatter plot of 10 hPa 30°N-90°N UARS MLS observations

709   and MERRA-2 sampled at the observation locations on the same day. The mean difference

710   is 0.08 ppmv (~1.3%) with the standard deviation of the differences of 0.48 ppmv and a

711   correlation coefficient of 0.94. This example demonstrates the ability of the reanalysis to

712   reproduce this dynamically driven structure accurately and corroborates the high

713   correlations seen in the northern hemisphere during winter (Figure 8i).

714

715   *6.4 Comparisons with MIPAS*

716

717   Figure 10 shows zonal mean and difference standard deviation statistics between MERRA-2

718   and MIPAS ozone as a function of latitude and pressure (60 hPa – 1 hPa), analogous to

719   Figure 8 for UARS MLS. The statistics are accumulated from 2003 – 2012, primarily in the

720   Aura time period. As in the case of UARS MLS comparisons, MERRA-2 is biased low with

721   respect to MIPAS in most of the stratosphere (Figure 10, top row). In particular, there is a

722   low bias of up to 8 % in the upper stratosphere between 3 hPa and 1 hPa. Between 30 hPa

723   and 10 hPa the negative bias is up to 3 % and up to 7 % below 30 hPa. The only region

724   where MERRA-2 is higher than MIPAS is the southern high latitudes in the upper

725   stratosphere in MAM and JJA (~1 %) and around 10 hPa in JJA (up to 5 %). Overall, the

726   low bias pattern is consistent throughout all seasons below 10 hPa.

727

728   The middle row of Figure 10 shows the MERRA-2 minus MIPAS standard deviations

729   (colors) and the standard deviations of the MIPAS ozone (contours) as a measure of ozone

730   variability. The results are very similar to those of the UARS MLS comparisons in the

731   previous section. In most of the stratosphere the difference standard deviation is less than 10

732   % and it is less than the variability, indicative of a good agreement between the reanalysis

733   and MIPAS data. Higher values of 15 % - 20 % are seen in high latitudes in winter and

734   spring when the variability is increased compared to other seasons.

735

736   The MIPAS – MERRA-2 correlations are shown in the bottom row of Figure 10 (panels i-l).

737   Overall, these are much higher than for UARS MLS (the bottom row of Figure 8), often in

738   excess of 0.8 in all seasons, especially in the extratropics. The weakest correlations (less

739   than 0.4) are found in the tropics where the ozone variability is the lowest in a given season,

740   consistent with the standard deviations shown in the middle row of Figure 10.

741

742   While the number of available MIPAS observations varies from year to year, there is

743   enough time continuity to calculate time series of monthly statistics. These are shown in

744   Figure 11 for MIPAS, MERRA and MERRA-2 at 9.42 hPa and 39.8 hPa in the 30°N-60°N

745   latitude band. The zonal mean time series show distinct annual cycles at both levels with

746   MERRA-2 biased low by ~ 0.6 ppmv at 9.42 hPa in summer during the MLS period and by

747   ~0.1 – 0.2 ppmv at 39.8 hPa. At 9.42 hPa MERRA is in a closer agreement with MIPAS

748　than MERRA-2, while the opposite is true at 39.4 hPa. The standard deviations of the

749　analysis minus MIPAS differences (panels (c) and (d)) also vary seasonally, with the

750　maximum differences in wintertime, as expected given the increased dynamical variability

751　in winter. At 9.42 hPa the standard deviations are similar for both reanalyses, ranging

752　between 0.25 ppmv in (~3 %) in summer and 0.6 ppmv (~10 %) during winter. At 39.4 hPa

753　the standard deviations are the same for both reanalyses in 2003, between 0.3 ppmv (~9 %,

754　summer) and 0.6 ppmv (15%, winter). However, in the Aura period (denoted by a yellow

755　vertical line in the figure), while the standard deviations against MERRA continue to

756　average ~ 0.4 ppmv, the standard deviations are much lower for MERRA-2, in the 1.8 to 3.2

757　ppmv range. The introduction of MLS observations led to much better agreement of the

758　analysis ozone with MIPAS data, consistent with the larger smoothing error at 40 hPa in

759　SBUV (Kramarova et al. 2012). Note the spike in the MERRA-2-MIPAS standard

760　deviations in April 2011 coinciding with an over three-week period of missing MLS

761　observations. The monthly MIPAS data counts are shown in panels (e) and (f) of Figure 11.

762

763　Comparisons in other latitude bands (not shown) reveal similar patterns: the reanalyses

764　capture the annual cycles of ozone at 9.42 hPa and 39.8 hPa and the agreement between

765　MERRA-2 and MIPAS improves in the Aura period, especially in the middle and high

766　latitudes. The greatest differences between the two reanalyses are found in the lower

767　stratosphere between 90°S and 60°S in austral winter when MERRA-2 performs much

768　better as a result of polar ozone observations from MLS.

769

770　*6.5 Summary of the comparisons against satellites*

771

772   This subsection summarizes comparisons of the MERRA-2 stratospheric ozone against

773   SAGE II, UARS MLS and MIPAS. The following are the main results of the zonal mean

774   comparisons.

775

776   • The mean MERRA-2 total ozone agrees with TOMS within 2 %. The difference

777   standard deviation does not exceed 6 %.

778   • In the upper stratosphere (3 hPa – 1 hPa) the reanalysis is biased low by 5 - 8 % with

779   respect to all three data sets in both SBUV and MLS periods.

780   • In the middle stratosphere there is agreement between the SAGE II and UARS MLS

781   comparisons in that, at least outside of high latitudes, MERRA-2 exhibits negative

782   bias of up to ~5 % between 30 hPa and 10 hPa and a small positive bias around 40

783   hPa. By contrast, MERRA-2 is low compared to MIPAS throughout the 40-10 hPa

784   layer. We note that the MIPAS comparisons are done mainly in the Aura period,

785   during which comparisons with SAGE II show almost no bias between 50 hPa and

786   10 hPa (Figure 7d-f). Hubert et al. (2016) find that MIPAS overestimates ozone

787   below 50 hPa compared to ozonesondes and lidar observations, especially in the

788   tropics where the differences exceed 10 % (see their Figure 6) and report a much

789   better agreement between SAGE II and independent data.

790   • Between 60 hPa and 50 hPa MERRA-2 is biased low compared to SAGE II, UARS

791   MLS and MIPAS observations..

792

793 The key conclusions regarding the stratospheric ozone variability in MERRA-2 are

794 summarized as follows.

795 • The difference standard deviations are within 20 % between 100 hPa and 1 hPa and

796 are much lower (within 10 %) throughout the middle stratosphere except in regions

797 of high ozone variability. Comparisons with the standard deviation of observed

798 ozone concentrations (Figures 8 and 10) indicate a very realistic representation of

799 variability in MERRA-2. This conclusion is corroborated by a case study showing

800 excellent agreement between the morphology of the MERRA-2 ozone field at 10 hPa

801 with UARS MLS during a complex polar transport event (Figure 9).

802 • The agreement between MERRA-2 and independent satellite data (SAGE II and

803 MIPAS) improves dramatically in the Aura period (from late 2004 onward),

804 particularly in the lower stratosphere where SBUV has reduced vertical resolution.

805

806 *6.6 Comparison with ozonesondes*

807 In this subsection we evaluate the MERRA-2 ozone in the upper troposphere and lower

808 stratosphere against balloon-borne ozonesonde observations. The comparisons are done as

809 follows: each ozonesonde profile is interpolated to a common vertical grid by averaging the

810 sonde measurements within a set of layers 1 km thick; the analysis profile the closest in time

811 and geolocation is mapped to the same vertical grid using cubic splines. In order to separate

812 tropospheric and stratospheric ozone the comparisons are done relative to the tropopause,

813 here defined as the 2 potential vorticity units (PVU) isosurface outside of the 10°S-10°N

814 latitude band and as the 100 hPa pressure surface within the band. The choice of 2 PVU was

815 motivated by the observation that the vertical gradient of ozone concentrations in the sonde

36

816   data changes sharply at this pressure. The same definition of the tropopause was employed

817   by Wargan et al. (2015). A tropopause-centered analysis was also used by Sofieva et al.

818   (2014) to construct an ozone climatology using ozonesonde and SAGE II data. The

819   advantage of this approach lies in its ability to separate the tropospheric and stratospheric

820   ozone content and emphasize the sharp cross-tropopause gradients in the ozone profiles. The

821   potential vorticity fields are taken from the MERRA-2 output. The results for 2003 (SBUV

822   period) and 2005 (Aura period) are shown in Figure 12. In both years all available sondes

823   were used. There were 1632 soundings in 2003 and 2425 in 2005. We note that the results

824   were almost unchanged when the 2005 data were randomly sub-sampled to a set of 1632

825   soundings. Note that the mean profiles are expressed in the units of partial pressure

826   (millipascals, mPa). As seen in panels (a) and (d) there is very good agreement between the

827   mean profiles from MERRA-2 and the ozonesondes in both years, however the vertical

828   gradient in the MERRA-2 ozone within the 2 km layer below the tropopause is sharper (and

829   in a better agreement with the sondes) in 2005. In the troposphere MERRA-2 is biased low

830   by up to ~0.2-0.3 mPa (5 – 12 %) in 2005. In 2003 there is a small high bias in the

831   uppermost troposphere and small low bias below the 5 km mark. In the stratosphere there

832   are patterns of small alternating biases not exceeding ~0.5 mPa (about 5 %). The standard

833   deviations of the sonde data shown as the magenta lines in panels (a) and (b) have a very

834   similar structure in both years. We note the enhanced variability within the 10 km layer

835   above the tropopause caused by vigorous advection associated with Rossby wave breaking

836   in the middle latitude surf zone during winter and spring (McIntyre and Palmer 1984). The

837   standard deviation of the MERRA-2 minus sonde differences (panels a, b, d, and e) is

838   smaller than the standard deviation of the sondes, especially above the tropopause,

839   indicative of a good representation of the ozone variability in that layer. This is particularly

840   evident in 2005 but it is also true in 2003. The ozonesonde – MERRA-2 correlations shown

841   in panels (c) and (f) are 0.7 – 0.8 below the tropopause in both years. In the layer of

842   maximum variability (within 10 km above the tropopause) the correlations are 0.8 and

843   above 0.9 in 2003 and 2005, respectively. These statistics are calculated from all available

844   ozonesondes and all seasons and therefore include latitudinal and seasonal as well as small-

845   scale, short-term variability. In order to test the robustness of the results we repeated the

846   comparisons for sonde locations between 0°E-60°E, 45°N-60°N in the spring (March-May)

847   2003 and 2005. The results are shown in Figure 13. The lower-stratospheric ozone

848   variability is larger here than in the global case (Figure 12) as is the gradient across the

849   tropopause. Both are captured very well by MERRA-2 in the Aura period (panels d and e)

850   but in the SBUV period the agreement is less good than in the global average (compare

851   panels a and b of Figures 12 and 13). The 'kink' in the ozonesonde profiles at 2 km and 5

852   km above the tropopause is a result of frequent isentropic intrusions of ozone-poor tropical

853   air into the mid-latitudinal lower stratosphere often associated with a double tropopause

854   (Sofieva et al. 2014). The feature is reproduced by MERRA-2 in both years but the

855   agreement with the sondes is closer in 2005. In 2003 the ozonesonde-analysis correlations

856   are low at the tropopause and the upper troposphere but they reach approximately 0.6 in the

857   layer between 5 km and 10 km above the tropopause and 0.8 above. In 2005 the

858   stratospheric correlations are higher, similar to the global case but they are much lower in

859   the troposphere except in the 2 km layer below the tropopause. The maximum ozone

860   variability relative to the mean occurs at the dynamical tropopause itself. This is also where

861   the reanalysis minus ozonesonde correlations reach a minimum as seen in Figures 12 and

862   14. The latter feature is also present but less pronounced when thresholds for the dynamical

863   tropopause of 4 PVU and 6 PVU are used.

864

865   Following Wargan et al. (2015) we perform separate comparisons of ozone integrated

866   within two layers: between 500 hPa and the tropopause (hereafter, upper troposphere or UT)

867   and between the tropopause and 50 hPa (lower stratosphere or LS). The results are

868   expressed in Dobson units. This analysis is motivated by the importance of an accurate

869   representation (including an assessment of errors) of the mid-latitude lower stratospheric

870   ozone and the separation of the stratospheric and upper-tropospheric air masses for studies

871   of tracer transport and radiative forcing. Monthly statistics calculated from MERRA and

872   MERRA-2 are compared against ozonesondes between 30°N and 60°N for the period 1991

873   to 2012. The results are shown in Figures 14 and 15 for the upper troposphere and lower

874   stratosphere, respectively. The MERRA-2 statistics are also summarized in Table 5. Note

875   that the number of sondes used per month increases in time from only a few before 1993 to

876   about 60 between 1997-2003 and then to 100 and above starting 2004 (Figure 14a). In

877   particular, the use of frequent soundings from field campaigns results in several spikes in

878   data counts. We first discuss the results in the UT shown in Figure 14. Both reanalyses

879   reproduce the month-to-month variations including the annual cycle seen in the sonde data

880   throughout the period of comparison. In the SBUV period the reanalyses are close to each

881   other with a small positive bias of no more than 3 DU with respect to the sondes. For

882   MERRA-2 the bias is within 10 % of the sonde mean (Table 5). In the Aura period

883   MERRA-2 becomes systematically lower than the sondes by 13.6 % on the average. The

884   bias has a seasonal dependence and varies between -1 DU and -4 DU. The existence of a

885   low bias in the UT in the Aura period is consistent with the results of Wargan et al. (2015)

886   where it was attributed to the absence of NO_x chemistry in the model and a low sensitivity

887   of OMI observations to the ozone below 500 hPa. The latter affects the UT through model

888   transport. Before the introduction of OMI and MLS observations in 2004 (and in MERRA)

889   the UT ozone is constrained by the partial column information from SBUV, which although

890   largely a priori information at these altitudes, prevents it from deviating significantly from a

891   climatological average. The standard deviation of the analysis – sonde differences (Figure

892   14c), similar for MERRA and MERRA-2, is generally between 2 DU and 4 DU. It is lower

893   for MERRA-2 in the Aura period by about 1 DU, indicative of a better representation of the

894   UT ozone variability after 2004. We performed a 95 % one-tailed F-test to assess the

895   significance of the ozonesonde-analysis standard deviations between the two reanalyses: if

896   MERRA-2 is closer to the sondes (the black line in Figure 14c is above zero) then the

897   statistic $\sigma\,(MERRA2 - sondes)/\sigma\,(MERRA - sondes)$ (where the symbol σ denotes the

898   standard deviation) is calculated and compared with the appropriate threshold given by the

899   F-distribution in order to determine if the difference is significant. If MERRA is closer to

900   the sondes then the same is applied to the reciprocal of the quotient above. Significant

901   differences are marked by blue (MERRA is closer to the sondes) and red (MERRA-2 is

902   closer) squares overplotted on the zero line. The sonde-MERRA-2 correlation also improves

903   from 0.81 to 0.9 between the two periods (Figure 14d).

904

905   In the LS the agreement between the reanalyses and sonde data is very close (Figure 15a).

906   For MERRA-2 the average bias is 3.8 % and 1.2 % in the SBUV and Aura periods,

907   respectively. The difference standard deviations are 11.2 % and 8.1% and the analysis-sonde

908   correlations are 0.96 and 0.98 (Figure 15b and c). In the Aura period MERRA-2 is

909   significantly closer to the sondes in terms of standard deviation of the analysis-sonde

910   differences as indicated by the difference line and the results of the F-test (the red square

911   marks in Figure 15b).

912

913   We reiterate that the structure of the tropospheric ozone profiles in MERRA-2 is mainly

914   controlled by transport by assimilated winds and parameterized convection. In the absence

915   of an explicit representation of NO_x chemistry and high-resolution observations in the

916   troposphere the quality of the ozone profiles below the tropopause is unavoidably degraded

917   compared to the stratosphere.

918

919   Figures 16 and 17 show comparisons between the UT and LS ozone from two reanalyses

920   and ozone sondes as a function of latitude for the SBUV (1991-2003) and Aura (2005-2012)

921   periods, respectively. The statistics are calculated within 5° wide latitude bins with the

922   condition that at least 10 soundings are available in a given bin. In the SBUV period the

923   mean agreement between the reanalyses and the ozonesondes in the LS is within 10 %

924   everywhere. It is within 20 % in the UT except between 30°S and the Equator, where both

925   reanalyses overestimate ozone by ~50% with respect to the sonde data. This large

926   overestimate results from a 5 DU positive bias in MERRA and MERRA-2 at three stations

927   located in the West Pacific ozone minimum: Samoa, Fiji and Tahiti. The minimum, a part of

928   the wave-one pattern in the tropical ozone, is associated with strong convection and

929   resulting lofting of ozone-poor air leading to a decrease of the tropospheric ozone column

930   (Thompson et al. 2003b). In the absence of accurate boundary layer ozone chemistry in the

931   model this feature is not correctly represented in the reanalyses. In the EOS Aura period the

932   agreement between MERRA-2 and ozonesondes in that region is much closer (Figure 17e).

933   Standard deviations of the differences (panels c and f) are between 10 % and 20 % in the LS

934   and between 15 % and 35 % in the UT. MERRA-2 exhibits a somewhat worse agreement

935   than MERRA in the northern high latitudes. In the Aura period (Figure 17) the LS bias is

936   within 10 % except in the tropics where the ozone columns are small (~ 20 DU) and where

937   MERRA-2 overestimates it by 20 %. This is consistent with the result seen in the profile

938   comparisons with SAGE II (Figure 7e) and the findings of Froidevaux et al. (2008) who

939   showed that version 2.2 of MLS ozone exhibit a positive bias of up to 50 % in the tropical

940   LS with respect to SAGE II data. A 20 % underestimate in the MERRA-2 UT in the

941   northern hemisphere (Figure 17e) has a similar latitudinal structure to that seen for the

942   reanalysis total ozone comparison with the MOD data in Figure 3. We attribute this

943   underestimate to the latitude-dependent bias in the OMI data (McPeters et al. 2015). As

944   expected, MERRA-2 performs better than MERRA in terms of the difference standard

945   deviations in the LS at almost all latitudes (Figure 17c). A smaller improvement is also seen

946   in the UT in that period (Figure 17f).

947

948   An example of the upper tropospheric (500 hPa to the tropopause) ozone distribution from

949   MERRA-2 are shown in Figure 18 for 1 July 2013. The plot illustrates the degree of spatial

950   variability resolved by the reanalysis and the realism of the large-scale patterns: the

951   summertime maximum in the northern hemisphere, a wave 1 pattern in the tropics with the

952   minimum over the Western Pacific and enhancements over the southern tropical Atlantic,

953   and high values along the subtropical jet stream over the southern Indian Ocean. Similar

954 seasonal features in tropospheric ozone fields have been reported by Ziemke et al. (2011)

955 who used an earlier version of MLS and OMI data to construct a monthly tropospheric

956 ozone climatology.

957

958 *6.7 Summary of the ozonesonde comparisons*

959

960 The main results of the comparisons with ozonesondes are summarized as follows

961

962 • In the lower stratosphere the agreement in terms of standard deviation in the

963 tropopause-50 hPa layer is within 20 % globally and about 10 % in the northern

964 middle latitudes. In the Aura period the agreement improves to less than 10 % in the

965 extratropics.

966 • The vertical structure of the LS ozone and the cross-tropopause gradient is much

967 better represented in the Aura period (Figures 12 and 13)

968 • In the upper troposphere the standard deviation of the reanalysis minus ozonesonde

969 differences is within 25 % and improves slightly in the Aura period but this

970 improvement comes at the expense of a low bias of 13.6 %.

971

972 **7. Conclusions**

973 In this study we described and evaluated the ozone product in GMAO's MERRA-2

974 reanalysis aiming to address the following two questions

975 1. How well does MERRA-2 represent the spatio-temporal variability of the

976 stratospheric and upper-tropospheric ozone?

977 2. What are the impacts of changing the ozone observing system in MERRA-2 in late

978 2004 when new types of data were introduced?

979

980 By combining observations from multiple sources with 6-hourly forecasts the data

981 assimilation methodology enables the production of high-frequency global gridded ozone

982 fields consistent with dynamical variability ranging from hourly to interannual scales, thus

983 providing an added value compared with ozone observations alone. The main goal of this

984 paper is to assess the usefulness of the MERRA-2 ozone for scientific applications by

985 highlighting the strengths as well as deficiencies of the reanalysis ozone.

986

987 The treatment of ozone data in MERRA-2 represents advancement over the original

988 MERRA reanalysis. In particular, the new reanalysis employs an improved version of

989 SBUV retrievals (version 8.6) between 1980 and 2004 and introduces EOS Aura data in the

990 later period with the OMI instrument providing over 300,000 total ozone observations

991 monthly and MLS constraining the stratosphere with observations at a relatively high-

992 resolution of about 2.5 km in the vertical and near-global day and night coverage.

993

994 The main findings are summarized as follows

995

996 • MERRA-2 total ozone agrees with TOMS data (1980-1993) very well, with less than

997 2 % bias and less than 6 % difference standard deviation, close to the assumed

998 observation error of 5 %.

44

999 • The bias in the stratospheric profiles with respect to SAGE II is reduced in 2005

1000 when MLS observations were assimilated compared to the SBUV period. The

1001 difference standard deviations with MIPAS are significantly smaller in the Aura

1002 period than in the SBUV period in the lower stratosphere. This is an expected

1003 consequence of the MLS profile data assimilation.

1004 • There is good representation of the variability of stratospheric ozone in MERRA-2.

1005 The difference standard deviations between the reanalysis and independent limb

1006 satellite data range from 11 % for SAGE II in the lower stratosphere to less than 5 %

1007 at 4.3 hPa. Comparisons against UARS MLS (the SBUV period) and MIPAS

1008 (mostly the Aura period) produced less that 10 % differences in most of the

1009 stratosphere except in regions of high variability such as winter high latitudes. The

1010 difference standard deviations do not exceed the standard deviation of the observed

1011 ozone concentrations (Figures 8 and 10), indicating the ability of the analysis to

1012 reproduce day-to-day ozone variability. As illustrated by a case study (Section 6.3.

1013 Figure 9) the reanalysis captures the complex dynamically driven morphology of the

1014 ozone field at 10 hPa during a cross-pole transport of middle-latitude high ozone in

1015 excellent agreement with UARS MLS data.

1016 • In the northern middle latitudes the UT ozone is 9.5 % high in the SBUV period and

1017 13.6 % low in the Aura period compared to ozonesondes. After 2004 the bias has

1018 some seasonal dependence but does not change much from year to year. By contrast,

1019 in the LS the bias is reduced in the Aura period but it is also small in the SBUV

1020 period (3.8 %). Overall, MERRA-2 reproduces the interannual to weekly

1021   (approximate frequency of ozonesonde observations) variability of the UTLS ozone

1022   as well as the sharp gradient of ozone concentrations across the tropopause.

1023

1024   The introduction of EOS Aura data leads to an improved representation of ozone in

1025   MERRA-2 but it also generates a discontinuity that should be taken into account in

1026   applications of the reanalysis to climate studies, in particular in trend analyses. The notable

1027   discontinuities in the ozone observing system are summarized in Table 6.

1028

1029   The results of this study provide a sound justification for using the MERRA-2 stratospheric

1030   and upper tropospheric ozone in scientific research, in particular for studies requiring high

1031   frequency, highly resolved global ozone maps with realistic variability consistent with

1032   dynamics. They highlight the value of NASA's research data from the EOS Aura mission to

1033   the reanalyses of long-term ozone records. While discontinuities have to be taken into

1034   account, MERRA-2 should be a viable dataset for use in monitoring the long-term changes

1035   in total and mid-stratospheric ozone as atmospheric halogen levels decrease. This includes

1036   the study of inter-annual changes (such as El Nino and the QBO), and potential impacts of

1037   extreme meteorological events (such as hurricanes) on ozone. For long-term time series

1038   analysis in the UTLS region extreme caution is recommended. Because of the larger SBUV

1039   smoothing error, signals in the UTLS that vary in altitude, such as the QBO discussed in

1040   section 6.3, will be represented differently in the SBUV and MLS time periods. Analyses

1041   requiring high vertical resolution, such as the study of inter-annual changes in mid-latitude

1042   stratosphere-troposphere exchange in relation to dynamical conditions, are recommended

1043   for the MLS-period only. In future reanalyses we will work to incorporate the SBUV

1044   averaging kernels to better represent the true vertical resolution of the input data. Despite

1045   these cautions, the results discussed here demonstrate progress towards representing ozone

1046   in a realistic fashion in Earth System reanalyses.

1047

1048   **Acknowledgments**

1058

1059   **References**

1060

1061   Bass, A. M. and Paur, R. J., 1985: The ultraviolet cross-sections of ozone, I, The

1062   measurements,in: Atmospheric Ozone, edited by: Zerefos, C. S. and Ghazi, A., D.

1063   Reidel, Norwell, Mass.

1064   Barath, F. T., M.C. Chavez, R.E. Cofield, D.A. Flower, M.A. Frerking, M.B. Gram, W.M.

1065   Harris, J.R. Holden, R.F. Jarnot, W.G. Kloezeman, G.J. Klose, G.K. Lau, M.S. Loo,

1066   B.J. Maddison, R.J. Mattauch, R.P. McKlnney, G.E. Peckham, H.M. Pickett, G. Siebes,

1067   F.S. Soltis, R.A. Suttie, J.A. Tarsala, J.W. Waters and W.J. Wilson, 1993: The Upper

1068   Atmosphere Research Satellite microwave limb sounder instrument. *J. Geophys. Res.*,

1069   98*(D6),* 10751–10762*, doi:*10.1029/93JD00798.

1070   Bhartia, P. K., R.D. McPeters, C.L. Mateer, L.E. Flynn, and C. Wellemeyer, 1996:

1071   Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet

1072   technique, J. Geophys. Res.-Atmos., 101, 18793–18806.

1073   Bhartia, P. K., and C. Wellemeyer, 2002: TOMS-V8 total O3 algorithm, in OMI Algorithm

1074   Theoretical Basis Document, vol. II, OMI Ozone Products, edited by P. K. Bhartia, pp.15–

1075   31,   NASA   Goddard   Space   Flight   Cent.,   Greenbelt,   Md.   (Available   at

1076   http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/index.php)

1077   Bhartia, P. K., R.D. McPeters, L.E. Flynn, S. Taylor, N.A. Kramarova, S. Frith, B. Fisher

1078   and M. DeLand, 2013: Solar Backscatter UV (SBUV) total ozone and profile algorithm.

1079   *Atmos. Meas. Tech.,* 6, 2533-2548, doi:10.5194/amt-6-2533-2013.

1080   Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu,

1081   Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliaev, Y.,

1082   Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O.,

1083   Schubert, S. D., and Suarez, M., 2015: MERRA-2: Initial Evaluation of the Climate.

1084   *NASA Tech. Rep. Series on Global Modeling and Data Assimilation,* NASA/TM–2015-

1085   104606, Vol. 43.

1086   Cohn, S.E., 1997: An Introduction to Estimation Theory. *J.Met.Soc of Japan,* Vol. 75, No.

1087   1B, pp. 257 – 288.

1088   Cortesi, U., J.C. Lambert, C. De Clercq , G. Bianchini, T. Blumenstock, A. Bracher, E.

1089   Castelli, V. Catoire, K.V. Chance, M. De Mazière, P. Demoulin, S. Godin-Beekmann,

1090   N. Jones, K. Jucks, C. Keim, T. Kerzenmacher, H. Kuellmann, J. Kuttippurath, M.

1091   Iarlori, G.Y. Liu, Y. Liu, I.S. McDermid, Y.J. Meijer, F. Mencaraglia, S. Mikuteit, H.

1092   Oelhaf, C. Piccolo, M. Pirre, P. Raspollini, F. Ravegnani, W.J. Reburn, G. Redaelli,

1093   J.J. Remedios, H. Sembhi, D. Smale, T. Steck, A. Taddei, C. Varotsos, C. Vigouroux,

1094   A. Waterfall, G. Wetzel and S. Wood, 2007: Geophysical validation of MIPAS-

1095   ENVISAT operational ozone data. *Atmos. Chem. Phys.,* 7, 4807-4867, doi:10.5194/acp-

1096   7-4807-2007.

1097   Coy L., K. Wargan, W.R. McCarty, S. Pawson, and A. Molod, 2015: Structure and

1098   Dynamics of the Quasi-Biennial Oscillation. *J. Climate,* **29**, 5339–5354, doi:

1099   10.1175/JCLI-D-15-0809.1.

1100   Damadeo, R. P., J.M. Zawodny, L.W. Thomason and N. Iyer, 2013: SAGE version 7.0

1101   algorithm: application to SAGE II. *Atmos. Meas. Tech.,* 6, 3539–3561,

1102   doi:10.5194/amt-6-3539-2013.

1103   Damadeo, R. P., J.M. Zawodny and L.W. Thomason, 2014: Reevaluation of stratospheric

1104   ozone trends from SAGE II data using a simultaneous temporal and spatial analysis.

1105   *Atmos. Chem. Phys.,* 14, 13455-13470, doi:10.5194/acp-14-13455-2014.

1106   Daumont, M., J. Brion, J. Charbonnier and J. Malicet, 1992: Ozone UV spectroscopy,

1107   I:Absorption cross-sections at room temperature. *J. Atmos. Chem*., 15, 145–155.

1108   DeLand, M. T., S.L. Taylor, L.K. Huang and B.L Fisher 2012: Calibration of the SBUV

1109   version 8.6 ozone data product, Atmos. Meas. Tech., 5, 2951-2967, doi:10.5194/amt-5-

1110   2951-2012.

1111   Dethof, A. and E.V. Hólm, 2004: Ozone assimilation in the ERA-40 reanalysis project.

1112   *Q.J.R. Meteorol. Soc.,* 130: 2851–2872. doi: 10.1256/qj.03.196

1113   Dragani R., 2011. On the quality of the ERA-Interim ozone reanalyses: comparisons with

1114   satellite data. *Q. J. R. Meteorol. Soc.* **137**: 1312–1326. DOI:10.1002/qj.821

1115   Fischer, H., M. Birk, C. Blom, B. Carli, M. Carlotti, T. von Clarmann, L. Delbouille, A.

1116   Dudhia, D. Ehhalt, M. Endemann, J.M. Flaud, R. Gessner, A. Kleinert, R. Koopman,

1117   J. Langen, M. López-Puertas, P. Mosner, H. Nett, H. Oelhaf, G. Perron, J. Remedios,

1118   M. Ridolfi, G. Stiller and R. Zander, 2008: MIPAS: an instrument for atmospheric and

1119   climate research. *Atmos. Chem. Phys.,* 8, 2151-2188, doi:10.5194/acp-8-2151-2008.

1120   Fleig, A. J., R. D. McPeters, P. K. Bhartia, B. M. Schlesinger, R. P. Cebula, K. F. Klenck, S.

1121   L. Taylor, and D. F. Heath, 1990: Nimbus 7 solar backscatter ultraviolet (SBUV) ozone

1122   products user's guide, Ref. Publ. 1234, NASA, Greenbelt, Md.

1123   Frith, S. M., N. A. Kramarova, R. S. Stolarski, R. D. McPeters, P. K. Bhartia, and G. J.

1124   Labow, 2014: Recent changes in total column ozone based on the SBUV Version 8.6

1125   Merged Ozone Data Set. *J. Geophys. Res. Atmos.,* 119, 9735-9751,

1126   doi:10.1002/2014JD021889.

1127   Froidevaux, L., et al., 2008: Validation of Aura Microwave Limb Sounder stratospheric

1128   ozone measurements. *J. Geophys. Res.,* 113, D15S20, doi:10.1029/2007JD008771.

1129   Global Modeling and Assimilation Office (GMAO), (2015a), MERRA-2 inst3_3d_asm_Nv:

1130   3D IAU State, Meteo- rology Instantaneous 3-hourly (0.625 3 0.5L72), version 5.12.4.

1131   Goddard Space Flight Center Distributed Active Archive Center, accessed June 2016,

1132   doi:10.5067/WWQSXQ8IVFW8.

1133   Global Modeling and Assimilation Office (GMAO), (2015b), MERRA-2 inst1_2d_asm_Nx:

1134   2d,3-Hourly,Instantaneous,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4,

1135   version 5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information

1136   Services Center (GES DISC),Accessed June 2016**,** doi:10.5067/3Z173KIE2TPD

1137   Herman, J. R., R. Hudson, R. McPeters, R. Stolarski, Z. Ahmad, X.-Y. Gu, S. Taylor, and C.

1138   Wellemeyer, 1991: A new self-calibration method applied to TOMS/SBUV

1139   backscattered ultraviolet data to determine long-term global ozone change. *J. Geophys.*

1140   *Res.,* 96, 7531-7545.

1141   Hilsenrath, E., R. P. Cebula, M. T. DeLand, K. Laamann, S. Taylor, C. Wellemeyer, and P.

1142   K. Bhartia, 1995: Calibration of the NOAA-11 Solar Backscatter UltraViolet (SBUV/2)

1143   ozone data set from 1989 to 1993 using in-flight calibration data and SSBUV, *J.*

1144   *Geophys. Res., 100*(D1), 1351-1366, doi:10.1029/94JD02611.

1145   Huang, L. K., R. P. Cebula, and E. Hilsenrath, 1998: New procedure for interpolating NIST

1146   FEL lamp irradiances, *Metrologia, 35*(4), 381-386, doi:10.1088/0026-1394/35/4/27.

1147   Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, J.-L. Baray, A.E.

1148   Bourassa, U. Cortesi, D.A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K.W.

1149   Hoppel, B.J. Johnson, E. Kyrölä, T. Leblanc, G. Lichtenberg, M. Marchand, C.T.

1150   McElroy, D. Murtagh, H. Nakane, T. Portafaix, R. Querel, J.M. Russell III, J. Salvador,

1151   H.G.J. Smit, K. Stebel, W. Steinbrecht, K.B. Strawbridge, R. Stübi, D.P.J. Swart, G.

1152   Taha, D.W. Tarasick, A.M. Thompson, J. Urban, J.A. E. van Gijsel, R. Van Malderen,

1153   P. von der Gathen, K.A. Walker, E. Wolfram and J.M. Zawodny, 2016: Ground-based

1154   assessment of the bias and long-term stability of 14 limb and occultation ozone profile

1155   data records, *Atmos. Meas. Tech.,* 9, 2497-2534, doi:10.5194/amt-9-2497-2016\

1156   Inness, A., F. Baier, A. Benedetti, I. Bouarar., S. Chabrillat, H. Clark, C. Clerbaux, P.

1157   Coheur, R.J. Engelen, Q. Errera, J. Flemming, M. George, C. Granier, J. Hadji-Lazaro,

1158   V. Huijnen, D. Hurtmans, L. Jones, J.W. Kaiser, J. Kapsomenakis, K. Lefever, J.

1159   Leitão, M. Razinger, A. Richter, M.G. Schultz, A.J. Simmons, M. Suttie, O. Stein, J.-

1160   Nn Thépaut, V. Thouret, M. Vrekoussis, C. Zerefos and the MACC team, 2013: The

1161   MACC reanalysis: an 8 yr data set of atmospheric composition. *Atmos. Chem. Phys.,* 13,

1162   4073–4109, doi:10.5194/acp-13-4073-2013.

1163   Inness, A., A.-M. Blechschmidt, I. Bouarar, S. Chabrillat, M. Crepulja, R.J. Engelen, H.

1164   Eskes, J. Flemming, A. Gaudel, F. Hendrick, V. Huijnen, L. Jones, J. Kapsomenakis,

1165   E. Katragkou, A. Keppens, B. Langerock, M. de Mazière, D. Melas, M. Parrington,

1166   V.H. Peuch, M. Razinger, A. Richter, M.G. Schultz, M. Suttie, V. Thouret, M.

1167   Vrekoussis, A. Wagner and C. Zerefos, 2015: Data assimilation of satellite-retrieved

1168   ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. *Atmos.*

1169   *Chem. Phys.,* 15, 5275-5303, doi:10.5194/acp-15-5275-2015.

1170   Janz, S., E. Hilsenrath, J. Butler, D.F. Heath, and R.P. Cebula, 1995: Uncertainties in

1171   radiance calibrations of backscatter ultraviolet (BUV) instruments. *Metrologia*, *32*(6),

1172   637-641, doi:10.1088/0026-1394/32/6/48.

1173   Kalnay, E. 2003: Atmospheric Modeling, Data Assimilation and Predictability.

1174   1063 Cambridge Univ. Press

1175   Kramarova, N., P. K. Bhartia, S. Frith, R. McPeters, and R. Stolarski, 2013: Interpreting

1176   SBUV smoothing errors: An example using the quasi-biennial oscillation, Atmos. Meas.

1177   Tech. Discuss., 6, 2721–2749, doi:10.5194/amtd-6-2721-2013.

1178   Kyrölä, E., M. Laine, V. Sofieva, J. Tamminen, S.-M. Päivärinta, S. Tukiainen, J.

1179   Zawodny and L. Thomason, 2013: Combined SAGE II–GOMOS ozone profile data set

1180   for 1984–2011 and trend analysis of the vertical distribution of ozone. *Atmos. Chem.*

1181   *Phys.,* 13, 10645-10658, doi:10.5194/acp-13-10645-2013.

1182   Labow, G. J., R. D. McPeters, P. K. Bhartia, *and* N. Kramarova, 2013: A comparison of 40

1183   years of SBUV measurements of column ozone with data from the Dobson/Brewer

1184   network, J. Geophys. Res. Atmos., 118, 7370–7378, *doi:*10.1002/jgrd.50503.

1185   Lahoz, W. A., Q. Errera, R. Swinbank and D. Fonteyn, 2007: Data assimilation of

1186   stratospheric constituents: a review. *Atmos. Chem. Phys.,* 7, 5745-5773,

1187   doi:10.5194/acp-7-5745-2007.

1188   Lahoz, W.A. and P. Schneider, 2014: Data Assimilation: Making sense of Earth

1189   observations. *Front. Environ. Sci.,* 28 May 2014, doi:10.3389/fenvs.2014.00016

1190   Levelt, P. F., G. H. J. van den Oord, M. R. Dobber, A. Mälkki, H. Visser, J. D. Vries, P.

1191   Stammes, J. O. V. Lundell, and H. Saari, 2006: The Ozone Monitoring Instrument, IEEE

1192   Trans. Geosci. Remote Sens., 44, 1093–1101, doi:10.1109/TGRS.2006.872333.

1193   Lefever, K., R. van der A, F. Baier, Y. Christophe, Q. Errera, H. Eskes, J. Flemming, A.

1194   Inness, L. Jones, J.-C. Lambert, B. Langerock, M.G. Schultz, O. Stein, A. Wagner and

1195   S. Chabrillat, 2015: Copernicus stratospheric ozone service, 2009–2012: validation,

1196   system intercomparison and roles of input data sets. *Atmos. Chem. Phys.,* 15, 2269-2293,

1197   doi:10.5194/acp-15-2269-2015.

1198   Livesey, N. J., W. G. Read, L. Froidevaux, J. W. Waters, M. L. Santee, H. C. Pumphrey, D.

1199   L. Wu, Z. Shippony, and R. F. Jarnot, 2003: The UARS Microwave Limb Sounder

1200   version 5 data set: Theory, characterization, and validation. *J. Geophys. Res.*, 108, *4378,*

1201   *doi:*10.1029/2002JD002273, D13.

1202 Livesey, N. J*., W. G. Read*, A. Lambert, R.E. Cofield, D.T.Cuddy, L. Froidevaux*, R.A.

1203 Fuller, R.F. Jarnot, J.H. Jiang, Y.B. Jiang, B.W. Knosp, L.J. Kovalenko, H.M. Pickett,

1204 H.C. Pumphrey, M.L. Santee, M.J. Schwartz, P.C. Stek, P.A. Wagner, J.W. Waters, and

1205 D.L. Wu, 2007: Version 2.2x Level 2 Data Quality and Description Document. JPL D-

1206 33509. Available from https://mls.jpl.nasa.gov/data/v2-2_data_quality_document.pdf

1207 Livesey, N. J*., W. G. Read*, P.A. Wagner, L. Froidevaux*, A. Lambert, G. Manney, L.F.M.

1208 Valle, H.C. Pumphrey, M.L. Santee, M.J. Schwartz, S. Wang, R.A. Fuller, R.F. Jarnot,

1209 B.W. Knosp and E. Martinez, 2015: Version 4.2x Level 2 Data Quality and Description

1210 Document. JPL D-33509 Rev. A. Available from https://mls.jpl.nasa.gov/data/v4-

1211 2_data_quality_document.pdf

1212 McCarty, W., L. Coy, R. Gelaro, D. Merkova, E. B. Smith, M. Sienkiewicz, and K. Wargan,

1213 2016: MERRA-2 Input Observations: Summary and Assessment. NASA/TM, *in*

1214 *preparation.*

1215 McIntyre, M. E., and T. N. Palmer, 1984: The 'surf zone' in the stratosphere. *J. Atmos. Terr.*

1216 *Phys.,***46,** 825–849.

1217 McPeters, R . D., and G. J. Labow, 1996: An assessment of the accuracy of 14.5 years of

1218 Nimbus 7 TOMS Version 7 ozone data by comparison with the Dobson Network,

1219 *Geophys. Res. Lett.,* 23, 3695-3698.

1220 McPeters, R. D., et al., 2008: Validation of the Aura Ozone Monitoring Instrument total

1221 column ozone product, *J. Geophys. Res.,* **113**, D15S14, doi:10.1029/2007JD008802.

1222 McPeters, R. D., Frith, S., and Labow, G. J., 2015: OMI total column ozone: extending the

1223   long-term data record. *Atmos. Meas. Tech.,* 8, 4845-4850, doi:10.5194/amt-8-4845-

1224   2015.

1225   Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5

1226   atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci.*

1227   *Model Dev.*, 8, 1339–1356.

1228   Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids.

1229   Journal of Computational Physics, 227 (1): 55-78 doi:10.1016/j.jcp.2007.07.022

1230   Raspollini, P., B. Carli, M. Carlotti, C. Ceccherini, A. Dehn, B.M. Dinelli, A. Dudhia, J.-

1231   M. Flaud, M. López-Puertas, F. Niro, J.J. Remedios, M. Ridolfi, H. Sembhi, L. Sgheri

1232   and T. von Clarmann, 2013: Ten years of MIPAS measurements with ESA Level 2

1233   processor V6 – Part 1: Retrieval algorithm and diagnostics of the products. *Atmos. Meas.*

1234   *Tech.,* 6, 2419-2439, doi:10.5194/amt-6-2419-2013.

1235   Sakazaki, T., M. Shiotani, M. Suzuki, D. Kinnison, J.M. Zawodny, M. McHug and K.A.

1236   Walker, 2015: Sunset–sunrise difference in solar occultation ozone measurements

1237   (SAGE II, HALOE, and ACE–FTS) and its relationship to tidal vertical winds. *Atmos.*

1238   *Chem. Phys.,* 15, 829-843, doi:10.5194/acp-15-829-2015.

1239   Sofieva, V. F., J. Tamminen, E. Kyrölä, T. Mielonen, P. Veefkind, B. Hassler and G.E.

1240   Bodeker, 2014: A novel tropopause-related climatology of ozone profiles, *Atmos. Chem.*

1241   *Phys.*, 14, 283-299, doi:10.5194/acp-14-283-2014.

1242 Stajner, I., et al., 2008: Assimilated ozone from EOS-Aura: Evaluation of the tropopause

1243 region and tropospheric columns. *J. Geophys. Res.*, 113, D16S32,

1244 doi:10.1029/2007JD008863.

1245 Takacs, L. L., M.J. Suárez and R. Todling, 2016: Maintaining atmospheric mass and water

1246 balance in reanalyses, Q J Roy Meteor Soc, 142, 1565-1573, 10.1002/qj.2763.

1247 Thompson, A. M., et al., 2003a: Southern Hemisphere Additional Ozonesondes (SHADOZ)

1248 1998–2000 tropical ozone climatology 1. Comparison with Total Ozone Mapping

1249 Spectrometer (TOMS) and ground-based measurements. *J. Geophys. Res.,* 108, 8238,

1250 doi:10.1029/2001JD000967, D2.

1251 Thompson A.M., et al. 2003b: Southern Hemisphere Additional Ozonesondes

1252 (SHADOZ) 1998-2000 tropical ozone climatology 2. Tropospheric variability and

1253 the zonal wave-one, J. Geophy. Res., 108, 8241, doi:10.1029/2002JD002241, 2003

1254

1255 Torres, O. and P.K. Bhartia, 1995: Effect of Stratospheric Aerosol on Ozone Profile from

1256 BUV Measurements. *Geophys. Res. Lett.,* 22, 235–238.

1257 Torres, O., J.R. Herman, P.K. Bhartia and Z. Ahmad, 1995: Properties of Mount-Pinatubo

1258 Aerosols as Derived from Nimbus-7 Total Ozone Mapping Spectrometer Measurements.

1259 *J. Geophys. Res.-Atmos.*, 100, 14043–14055.

1260 Vasilkov, A.P., J. Joiner, K. Yang, and P. K. Bhartia, 2004: Improving total column ozone

1261 retrievals by using cloud pressures derived from Raman scattering in the UV.

1262 *Geophysical Research Letters*, 31, L20109, doi:10.1029/2004GL020603.

1263 Wang, H. J., D. M. Cunnold, L. W. Thomason, J. M. Zawodny, and G. E. Bodeker, 2002:

1264   Assessment of SAGE version 6.1 ozone data quality. *J. Geophys. Res*., 107(D23), 4691,

1265   doi:10.1029/2002JD002418.

1266 Wargan, K., S. Pawson, M. A. Olsen, J. C. Witte, A. R. Douglass, J. R. Ziemke, S. E.

1267   Strahan, and J. E. Nielsen, 2015: The global structure of upper troposphere- lower

1268   stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data. *J. Geophys.*

1269   *Res. Atmos.,* 120, 2013–2036, doi:10.1002/ 2014JD022493.

1270 Waters, J. W. and Co-authors, 2006: The Earth Observing System Microwave Limb

1271   Sounder (EOS MLS) on the Aura satellite. *IEEE Trans. Geosci. Remote Sens*., **44**,

1272   1075–1092.

1273 Wellemeyer, C.G., P. K. Bhartia, S. L. Taylor, W. Qin, and C. Ahn, 2004: Version 8 Total

1274   Ozone Mapping Spectrometer (TOMS) Algorithm, paper presented at Quadrennial Ozone

1275   Symposium, Eur. Comm., Kos, Greece.

1276 World Meteorological Organization (WMO 2014): Scientific Assessment of Ozone

1277   Depletion: 2014. Global Ozone Research and Monitoring Project – Report no. 55

1278 Ziemke, J. R., S. Chandra, G.J. Labow, P.K. Bhartia, L. Froidevaux and J.C. Witte, 2011: A

1279   global climatology of tropospheric and stratospheric ozone derived from Aura OMI and

1280   MLS measurements*. Atmos. Chem. Phys.,* 11, 9237-9251, doi:10.5194/acp-11-9237-

1281   2011.

1282 Ziemke, J. R., et al., 2014: Assessment and applications of NASA ozone data products

1283   derived from Aura OMI/MLS satellite measurements in context of the GMI chemical

1284   transport model, J. Geophys. Res. Atmos., 119, doi:10.1002/2013JD020914.

1285

1286

1287

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

1302

1303

1304

1305

1306   **Tables**

1307

1308   Table 1. Ozone data assimilated in MERRA-2

Instrument and platform	Time period	Type of measurement	Coverage
SBUV Nimbus-7	1 Jan 1980 – 31 May 1990	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-11	1 Jan 1989 – 4 Feb 1995	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-14	5 Feb 1995 – 30 Jun 2001	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-16	1 Jan 2001 – 30 Sep 2004	Partial and total columns	Sunlit atmosphere, surface to TOA
SBUV/2 NOAA-17	17 Nov 2002 – 30 Sep 2004	Partial and total columns	Sunlit atmosphere, surface to TOA
OMI EOS Aura	1 Oct 2004 - present	Total column	Sunlit atmosphere
MLS EOS Aura	1 Oct 2004 - present	Mixing ratio profiles	82°S – 82°N mainly stratosphere

1309

1310

1311

1312

1313

1314   Table 2. Statistical comparisons of the MERRA-2 total ozone against TOMS data. The

1315   percent values are relative to TOMS averages.

	Bias		Error standard deviation	
	[DU]	[%]	[DU]	[%]
90°S-60°S	1.80	0.63	15.15	5.32
60°S-30°S	-4.49	-1.43	10.84	3.46
30°S-30°N	-4.92	-1.83	7.44	2.77
30°N-60°N	-2.60	-0.77	12.58	3.72
60°N-90°N	5.09	1.45	19.35	5.53

1316

1317

1318

1319

1320

1321

1322

1323

1324

1325

1326

1327

1328

1329   Table 3. Statistical comparisons of the MERRA-2 stratospheric ozone against SAGE II data

1330   south of 30°S. Results obtained using sunset only data are shown in parentheses.   The

1331   percent values are relative to SAGE II averages.

1332

	Bias		Error standard deviation	
	[ppmv]	[%]	[ppmv]	[%]
4.3 hPa	0.14 (0.06)	2 (0.9)	0.31 (0.3)	4.6 (4.3)
10.1 hPa	0.09 (0.07)	1.5 (1.1)	0.4 (0.42)	6.4 (6.5)
42.7 hPa	0.06 (0.07)	1.9 (2)	0.35 (0.35)	11 (11)

1333

1334

1335

1336

1337

1338

1339

1340

1341

1342

1343

1344

1345   Table 4. Statistical comparisons of the MERRA-2 stratospheric ozone against SAGE II data

1346   north of 30°N. Results obtained using sunset only data are shown in parentheses. The

1347   percent values are relative to SAGE II averages.

	Bias		Error standard deviation	
	[ppmv]	[%]	[ppmv]	[%]
4.3 hPa	0.07 (0.02)	1.1 (0.3)	0.32 (0.31)	4.6 (4.5)
10.1 hPa	0.08 (0.06)	1.3 (1.0)	0.37 (0.37)	5.7 (5.8)
42.7 hPa	0.05 (0.05)	1.6 (1.6)	0.37 (0.37)	11.3 (11.3)

1348

1349

1350

1351

1352

1353

1354

1355

1356

1357

1358

1359

1360

1361   Table 5. Statistical comparisons of the MERRA-2 UT and LS ozone against ozonesonde

1362   measurements between 30°N and 60°N. The percent values are relative to ozonesondes

1363   averages.

	Bias				Error standard deviation				Correlation	
	SBUV period		Aura period		SBUV period		Aura period		SBUV period	Aura period
	[DU]	[%]	[DU]	[%]	[DU]	[%]	[DU]	[%]		
Upper troposphere	1.22	9.5	-1.77	-13.6	3.15	24.5	2.74	21.1	0.81	0.9
Lower stratosphere	3.65	3.8	1.23	1.2	10.62	11.2	8.27	8.1	0.96	0.98

1364

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374   Table 6. Discontinuities in the ozone observing system.

Time / period	Source of discontinuity
1-31 March 1991	Missing NOAA-11 SBUV observations; no ozone data are assimilated
Late 1994	SBUV coverage limited to latitudes north of 30°S
1 October 2004	Introduction of EOS Aura data; SBUV data turned off
27 March - 18 April 2011	MLS data outage
1 June 2015	Transition from Version 2.2 to Version 4.2 of MLS retrievals
1 May 2016	The 261 hPa MLS level turned off
29 May - 13 June 2016	OMI data outage (not discussed in this study but included here for completeness)

1375

1376

1377

1378

Figure captions

1380 **Figure 1.** The total column ozone observations assimilated in MERRA-2. (a) the monthly
1381 averaged number of total ozone observations per day from SBUV and OMI instruments as a
1382 function of time. The SBUV and Aura periods are separated by a thick vertical divider. Note
1383 that the daily numbers for SBUV and OMI are given in the units of 10,000 and 100,000,
1384 respectively. (b) the monthly latitudinal coverage of the total ozone observations. Colors
1385 denote different instruments as indicated in the legend in panel b.
1386

1387 **Figure 2.** Comparisons of the monthly statistics total ozone from MERRA (blue) and
1388 MERRA-2 (red) with TOMS data in 5 latitude bands: (a, b) 90°S-60°S, (c, d) 60°S-30°S, (e,
1389 f) 30°S-30°N, (g, h) 30°N-60°N, and (i, j) 60°N-90°N. The left column (panels a, c, e, g and
1390 i) shows time series of reanalysis minus TOMS differences expressed as percentages of
1391 TOMS monthly averages. The right panels (b, d, f, h, and j) plot the corresponding standard
1392 deviations of reanalysis minus TOMS differences relative to TOMS. The times of El
1393 Chichón (March-April 1982) and Mt. Pinatubo (June 1991) eruptions are marked by yellow
1394 dashed lines in the central panels (e and f). The only full month of missing ozone data
1395 (March 1991) is indicated by solid yellow lines in the same panels.
1396

1397 **Figure 3.** Comparisons of the MERRA-2 total ozone against the Merged Ozone Data Set
1398 (MOD). The plot shows the reanalysis minus MOD difference as a percentage of MOD for
1399 two periods: 1980-2003 (black) and 2005-2014 (blue).
1400

1401 **Figure 4.** (a) The 1991-2015 time series of total ozone in Dobson units at the South Pole
1402 derived from individual ozonesonde measurements (black) and collocated reanalysis:
1403 MERRA (blue) and MERRA-2 (red). (b) reanalysis minus ozonesonde differences divided
1404 by sonde total ozone and expressed in percent: MERRA (blue) and MERRA-2 (red). The
1405 SBUV and Aura periods are separated by the black vertical line.
1406

1407 **Figure 5.** Time series of annually averaged ozone differences integrated between 208 hPa
1408 and 0.2 hPa: MERRA minus SAGE II (blue) and MERRA-2 minus SAGE II (red). The
1409 differences are expressed as percent of the SAGE II annual means. The averaging is done in
1410 three latitude bands: south of 30°S (a), between 30°S and 30°N (b) and north of 30°N (c).
1411 The years 1992 and 1993 are omitted because of a lower quality of SAGE II data in the
1412 lower stratosphere following the Mt. Pinatubo eruption.
1413

1414 **Figure 6.** Time series of annually averaged statistics from SAGE II, MERRA and MERRA-
1415 2 south of 30°S at 42.7 hPa (a, d and g), 10.1 hPa (b, e and h) and 4.3 hPa (c, f and i). The
1416 top row (a-c): the annual mean mixing ratio from SAGE II (black), MERRA (blue) and
1417 MERRA-2 (red). The middle row (d-f): the standard deviation of the MERRA minus SAGE
1418 II (blue) and MERRA-2 minus SAGE II (red) differences. The total annual SAGE II data
1419 counts for each level are shown in the bottom row (g-i). The units are parts per million by
1420 volume
1421

1422 **Figure 7.** Profiles of MERRA – SAGE II (blue) and MERRA-2 – SAGE II (red)
1423 comparisons relative to the SAGE II average calculated for all available data in January-

1424 August 2003 (a-c) and 2005 (d-f) and expressed in percent. Statistics for three latitude bands
1425 are shown: south of 30°S (a and d), between 30°S and 30°N (b and e) and north of 30°N (c
1426 and f). The solid lines represent the mean differences between the reanalysis and SAGE II.
1427 The standard deviations of the differences are shown as dashed lines. The SAGE II
1428 observation levels are indicated by dots superimposed on the solid lines. All statistics are
1429 shown as a percentage of the SAGE II mean over 2003 (top row) or 2005 (bottom).
1430
1431 **Figure 8**. Top row: the zonal mean difference between MERRA-2 and UARS MLS ozone
1432 computed from all UARS MLS data between 1991 and 1996 for December – February (a),
1433 March-May (b), June-August (c) and September-November (d). The difference is expressed
1434 as a percent of the zonal mean UARS MLS ozone. Middle row: The standard deviation of
1435 the difference between the two data sets, relative to the zonal mean UARS MLS
1436 observations (colors) and relative standard deviation of UARS MLS ozone (white contours)
1437 for December – February (e), March-May (f), June-August (g) and September-November
1438 (h). Bottom row: correlations between MERRA-2 and UARS MLS for December –
1439 February (i), March-May (j), June-August (k) and September-November (l). The black
1440 contours are 0.5, 0.7 and 0.9, the 0.7 contour is thickened. The comparisons are done in the
1441 57 hPa – 1 hPa range.
1442
1443 **Figure 9**. (a) Color background: MERRA-2 ozone between 30°N and 90°N at 10 hPa on 21
1444 December 1993 at 6Z. UARS MLS observations made between 1:30Z and 10:30Z are
1445 shown as filled circles color-coded by the retrieved mixing ratios. The color contour spacing
1446 is 0.2 ppmv. The highest latitude observed by SBUV around the same day (51°N) is marked
1447 by black dots. (b) a scatter plot of all UARS MLS observations at 10 hPa on 21 December
1448 and collocated MERRA-2 ozone the closest in time to the MLS data (no more than 1.5 h
1449 apart). The diagonal is shown as a dotted line and the solid black line is the linear fit. The
1450 units are parts per million by volume.
1451
1452 **Figure 10**. Top row: the zonal mean difference between MERRA-2 and MIPAS ozone
1453 computed from all MIPAS data between 2003 and 2012 for December – February (a),
1454 March-May (b), June-August (c) and September-November (d). The difference is expressed
1455 as a percent of the zonal mean MIPAS ozone. Middle row: The standard deviation of the
1456 difference between the two data sets, relative to the zonal mean MIPAS observations
1457 (colors) and relative standard deviation of MIPAS ozone (white contours) for December –
1458 February (e), March-May (f), June-August (g) and September-November (h). Bottom row:
1459 Correlations between MERRA-2 and MIPAS data for December – February (i), March-May
1460 (j), June-August (k) and September-November (l). The black contours are 0.5, 0.7 and 0.9,
1461 the 0.7 contour is thickened. The comparisons are done in the 57 hPa – 1 hPa range.
1462
1463 **Figure 11.** Time series of MIPAS, MERRA, and MERRA-2 monthly ozone statistics. 30°N
1464 – 60°N. Top row: the monthly average from MIPAS (black), MERRA (blue) and MERRA-2
1465 (red) at 9.42 hPa (a) and 39.8 hPa (b). Middle row: the reanalysis-MIPAS difference
1466 standard deviation at 9.42 hPa (c) and 39.8 hPa (d). The units are parts per million by
1467 volume. Bottom row: The corresponding monthly MIPAS observation counts at 9.42 hPa (e)
1468 and 39.8 hPa (f). The yellow lines in panels a-d mark the transition from the SBUV to EOS
1469 Aura period in MERRA-2.

1470

1471 **Figure 12**. Statistical comparisons of MERRA-2 ozone against ozonesondes available in
1472 2003 (a-c) and 2005 (d-f). All data are interpolated to the tropopause-based vertical
1473 coordinate. (a and d) the average ozone partial pressure from ozone sondes (black) and
1474 MERRA-2 (red), the standard deviation of ozonesonde measurements (magenta), and the
1475 standard deviation of MERRA-2 minus ozonesonde differences (shaded). (b and e) The
1476 mean MERRA-2 minus ozonesonde differences (black line) and the standard deviation of
1477 the difference (shaded), all expressed as a percent of sonde ozone. (c and f) the MERRA-2 –
1478 ozonesonde correlations. The filled circles are the levels to which all data are interpolated
1479 and the dashed line is the dynamical tropopause. The units in a and d are millipascals.

1480

1481 **Figure 13.** The same as Figure 13 but with soundings limited to 0°E-60°E, 45°N-60°N,
1482 March-May.

1483

1484 **Figure 14**. Time series of various statistics comparing ozonesondes between 30°N and 60°N
1485 (black), MERRA (blue) and MERRA-2 (red) integrated in the upper troposphere (500 hPa
1486 to the dynamical tropopause). (a) number of sondes in each month. (b) the monthly averaged
1487 ozone (solid lines with dots) and the reanalysis minus sonde differences in Dobson units
1488 (dotted lines). (c) the standard deviation of the reanalysis minus ozonesonde differences
1489 (blue and red for MERRA and MERRA-2, respectively) and their difference (MERRA-2
1490 standard deviation, minus MERRA standard deviation,, black) in Dobson units. The red
1491 (blue) squares on the zero line indicate months when MERRA-2 (MERRA) is closer to the
1492 sondes than MERRA (MERRA-2) at the 95% significance level. (d) monthly reanalysis-
1493 ozonesonde correlations. The grey bars indicate the months for which the correlations are
1494 not statistically significant at 95%. The vertical yellow lines indicate October 2004 when
1495 Aura data were introduced in MERRA-2.

1496

1497 **Figure 15**. Time series of various statistics comparing ozonesondes between 30°N and 60°N
1498 (black), MERRA (blue) and MERRA-2 (red) integrated in the lower stratosphere (between
1499 the dynamical tropopause and 50 hPa). (a) the monthly averaged ozone (solid lines with
1500 dots) and the reanalysis minus sonde differences in Dobson units (dotted lines). (b) the
1501 standard deviation of the reanalysis minus ozonesonde differences (blue and red for
1502 MERRA and MERRA-2, respectively) and their difference (MERRA-2 standard deviation.
1503 minus MERRA standard deviation, black) in Dobson units. The red (blue) squares on the
1504 zero line indicate months when MERRA-2 (MERRA) is closer to the sondes than MERRA
1505 (MERRA-2) at the 95% significance level. (c) monthly reanalysis-ozonesonde correlations.
1506 The vertical yellow lines indicate October 2004 when Aura data were introduced in
1507 MERRA-2.

1508

1509 **Figure 16.** Comparisons of LS (a-c) and UT (d-f) ozone from MERRA (blue) and MERRA-
1510 2 (red) with ozonesondes (black) as a function of latitude. All available sondes between
1511 1991 and 2003 are used. The data are binned into 36 latitude bins, each 5° wide. Shown are
1512 the average in Dobson units (a and d), the reanalysis minus sonde difference relative to the
1513 sonde mean in percent (b and e) and the standard deviation of the difference also relative to
1514 the sonde average in percent (c and f).

1515

1516 **Figure 17.** As in Figure 16 but for the 2005-2012 period.
1517
1518 **Figure 18**. Mean ozone mixing ratio in parts per million by volume between 500 hPa and
1519 the tropopause from MERRA-2 on 1 July 2013. The areas where the tropopause lies below
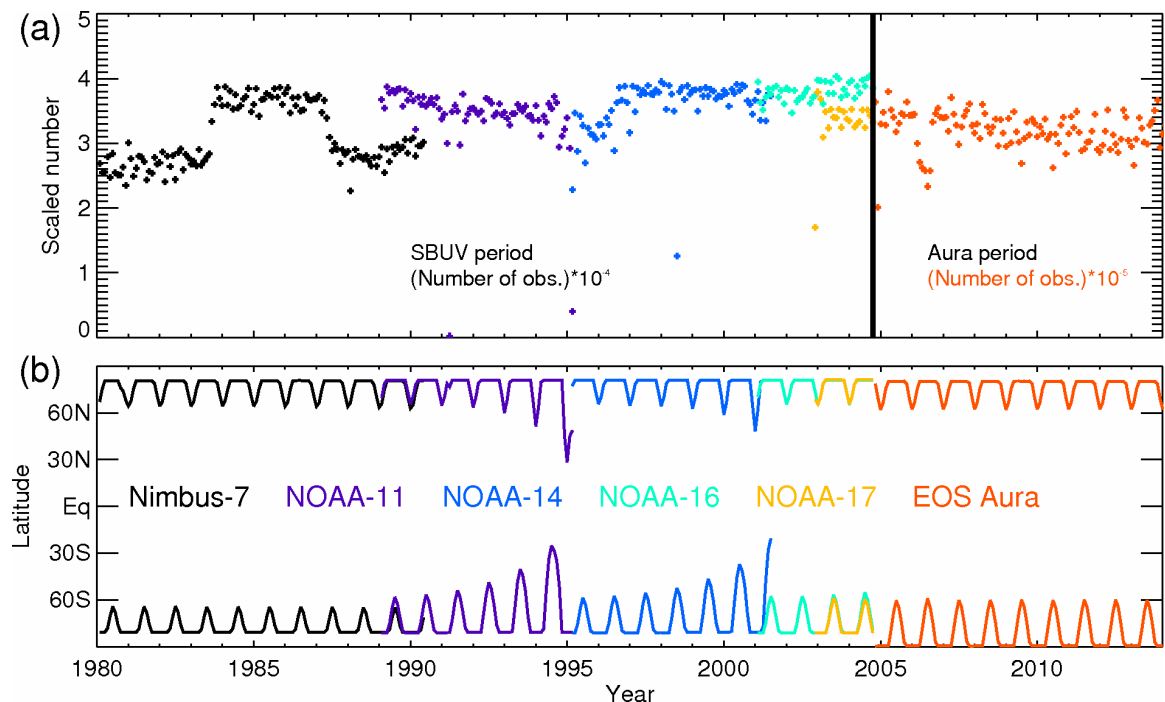1520 the 500-hPa surface are marked white.
1521
1522
1523

1524

1525

1526

1527

1528

1529

1530

1531

1532

1533

1534

1535

1536

1537

1538

1539

1540

1541

1542

Figures

1546   ***Figure 1.*** *The total column ozone observations assimilated in MERRA-2. (a) the monthly*
1547   *averaged number of total ozone observations per day from SBUV and OMI instruments as*
1548   *a function of time. The SBUV and Aura periods are separated by a thick vertical divider.*
1549   *Note that the daily numbers for SBUV and OMI are given in the units of 10,000 and*
1550   *100,000, respectively. (b) the monthly latitudinal coverage of the total ozone*
1551   *observations. Colors denote different instruments as indicated in the legend in panel b.*
1552
1553
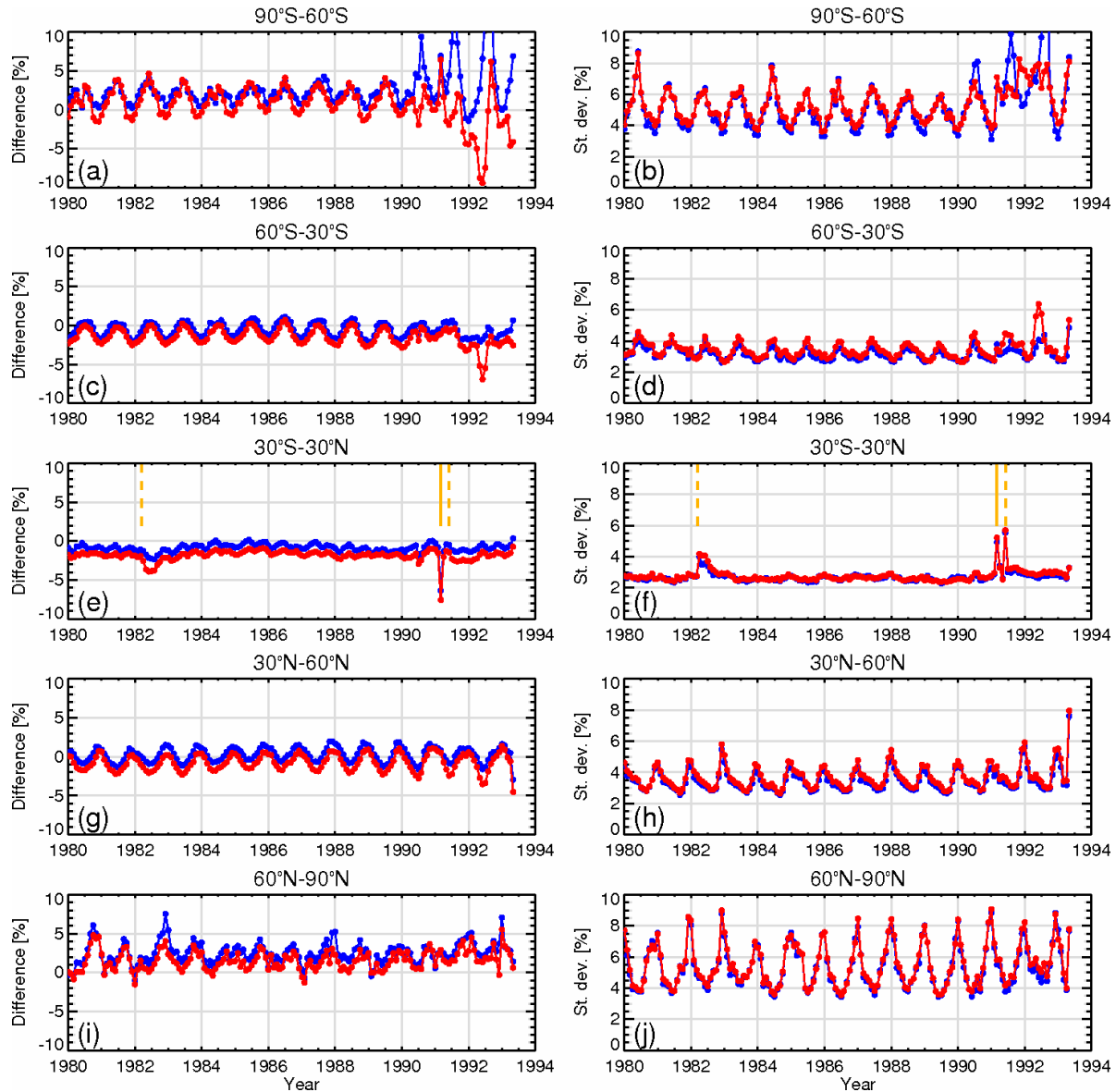
1554
1555 **_Figure 2._** _Comparisons of the monthly statistics total ozone from MERRA (blue) and_
1556 _MERRA-2 (red) with TOMS data in 5 latitude bands: (a, b) 90°S-60°S, (c, d) 60°S-30°S, (e,_
1557 _f) 30°S-30°N, (g, h) 30°N-60°N, and (i, j) 60°N-90°N. The left column (panels a, c, e, g and_
1558 _i) shows time series of reanalysis minus TOMS differences expressed as percentages of_
1559 _TOMS monthly averages. The right panels (b, d, f, h, and j) plot the corresponding_
1560 _standard deviations of reanalysis minus TOMS differences relative to TOMS. The times of_
1561 _El Chichón (March-April 1982) and Mt. Pinatubo (June 1991) eruptions are marked by_
1562 _yellow dashed lines in the central panels (e and f). The only full month of missing ozone_
1563 _data (March 1991) is indicated by solid yellow lines in the same panels._
1564

71
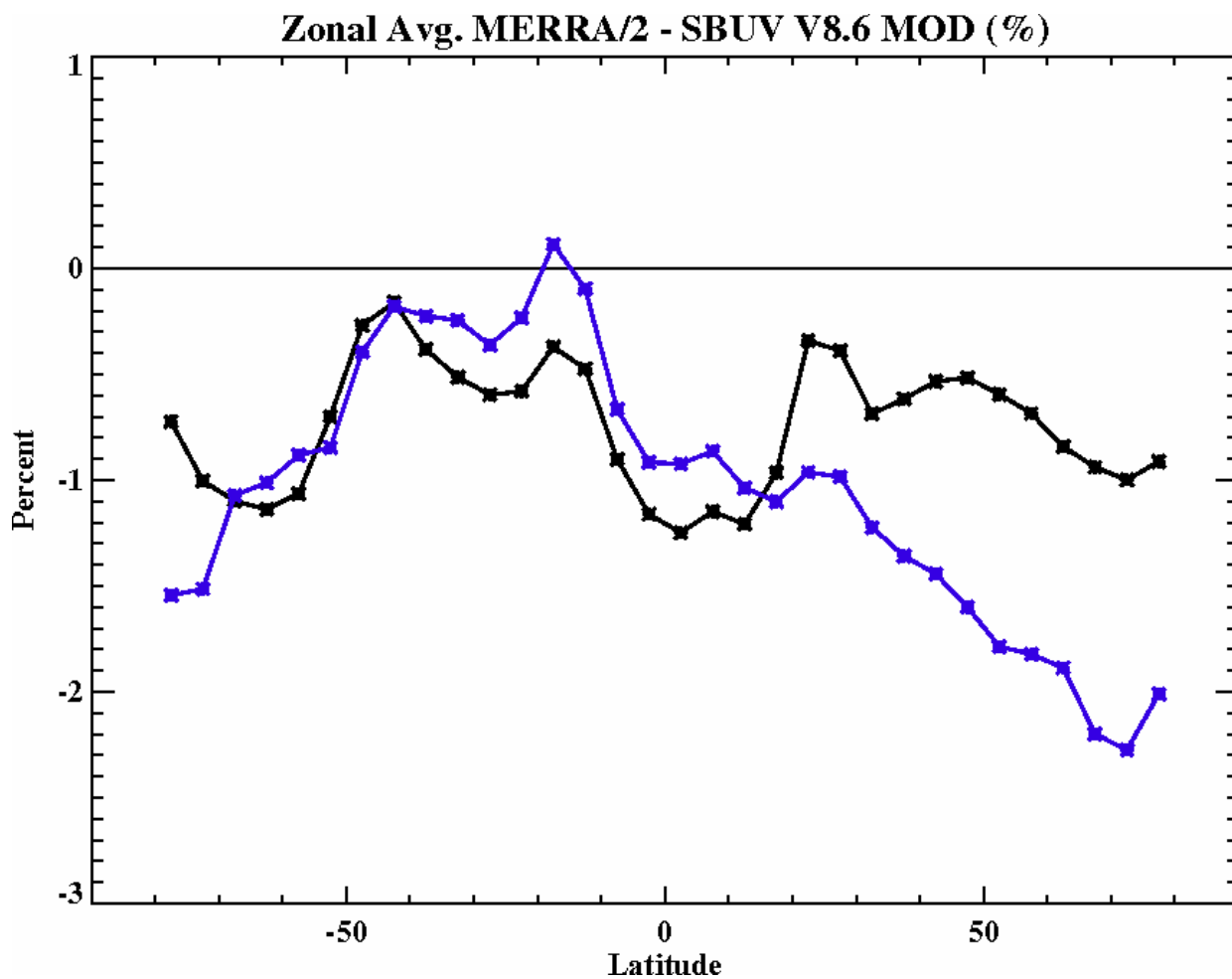
Figure 3. *Comparisons of the MERRA-2 total ozone against the Merged Ozone Data Set (MOD). The plot shows the reanalysis minus MOD difference as a percentage of MOD for two periods: 1980-2003 (black) and 2005-2014 (blue).*
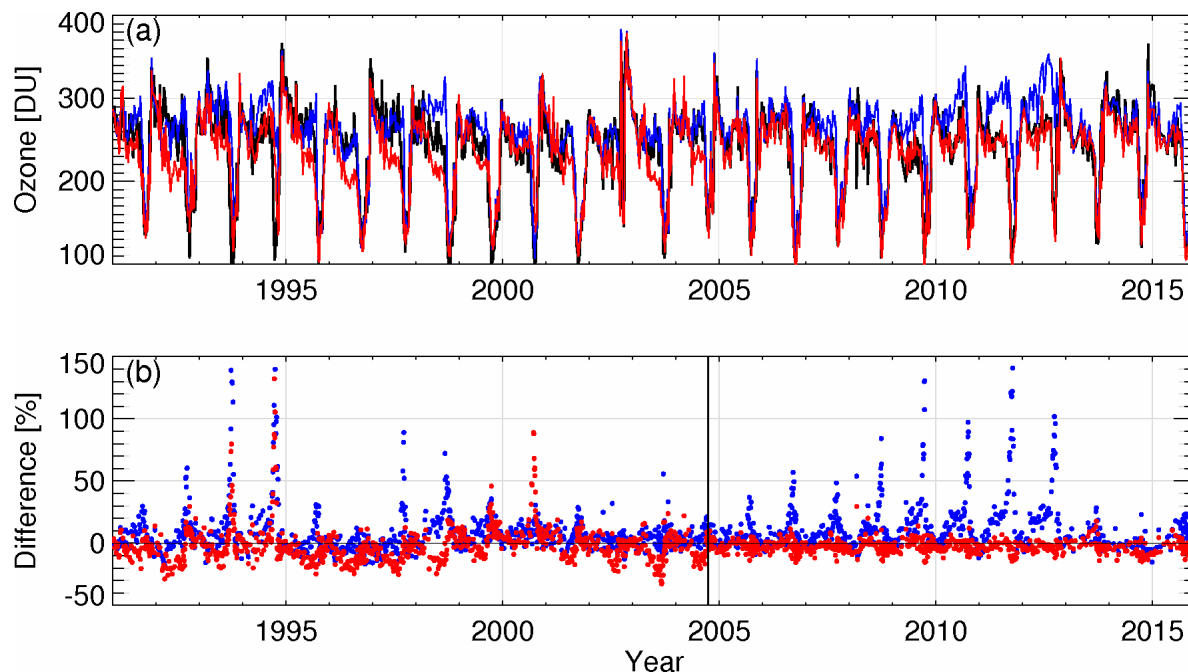
Figure 4. *(a) The 1991-2015 time series of total ozone in Dobson units at the South Pole derived from individual ozonesonde measurements (black) and collocated reanalysis: MERRA (blue) and MERRA-2 (red). (b) reanalysis minus ozonesonde differences divided by sonde total ozone and expressed in percent: MERRA (blue) and MERRA-2 (red). The SBUV and Aura periods are separated by the black vertical line.*
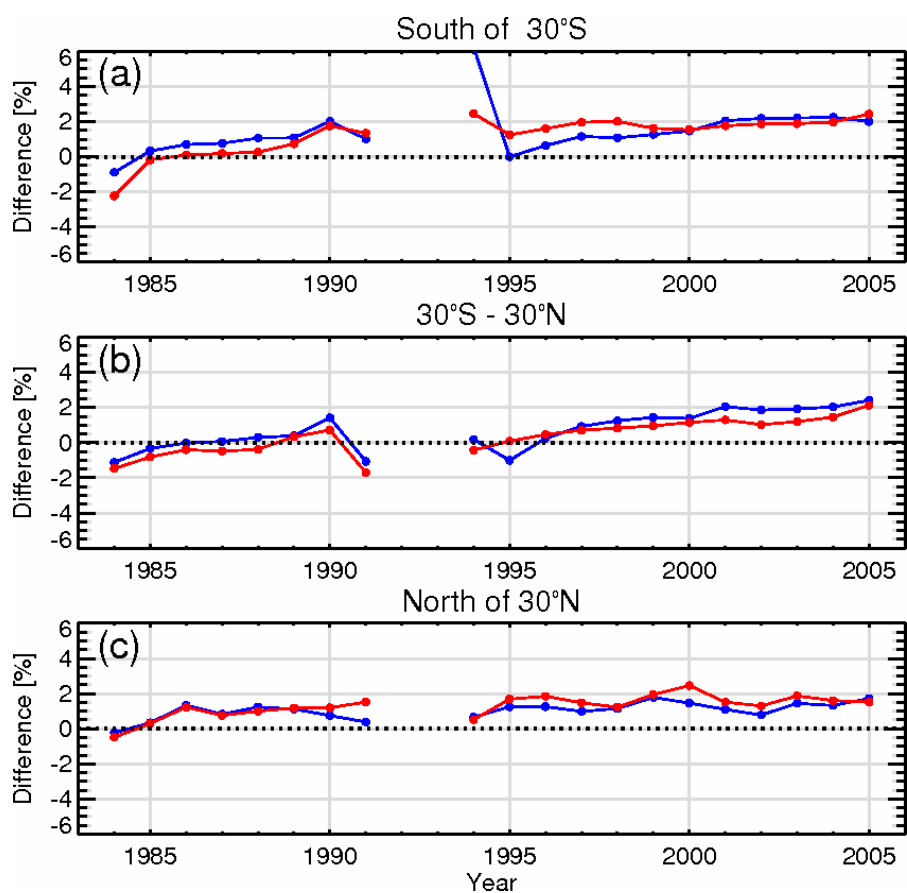
1577
1578
1579 **Figure 5.** *Time series of annually averaged ozone differences integrated between 208*
1580 *hPa and 0.2 hPa: MERRA minus SAGE II (blue) and MERRA-2 minus SAGE II (red). The*
1581 *differences are expressed as percent of the SAGE II annual means. The averaging is done*
1582 *in three latitude bands: south of 30°S (a), between 30°S and 30°N (b) and north of 30°N*
1583 *(c). The years 1992 and 1993 are omitted because of a lower quality of SAGE II data in*
1584 *the lower stratosphere following the Mt. Pinatubo eruption.*
1585
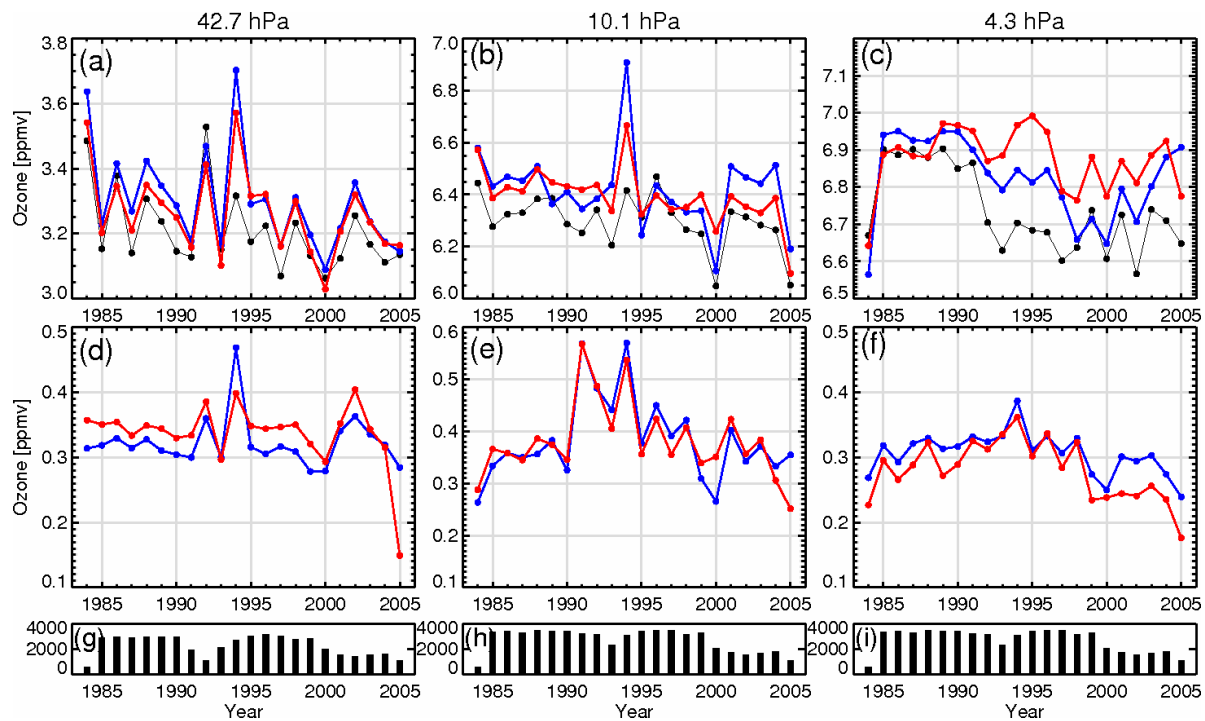
Figure 6. *Time series of annually averaged statistics from SAGE II, MERRA and MERRA-2 south of 30°S at 42.7 hPa (a, d and g), 10.1 hPa (b, e and h) and 4.3 hPa (c, f and i). The top row (a-c): the annual mean mixing ratio from SAGE II (black), MERRA (blue) and MERRA-2 (red). The middle row (d-f): the standard deviation of the MERRA minus SAGE II (blue) and MERRA-2 minus SAGE II (red) differences. The total annual SAGE II data counts for each level are shown in the bottom row (g-i). The units are parts per million by volume*
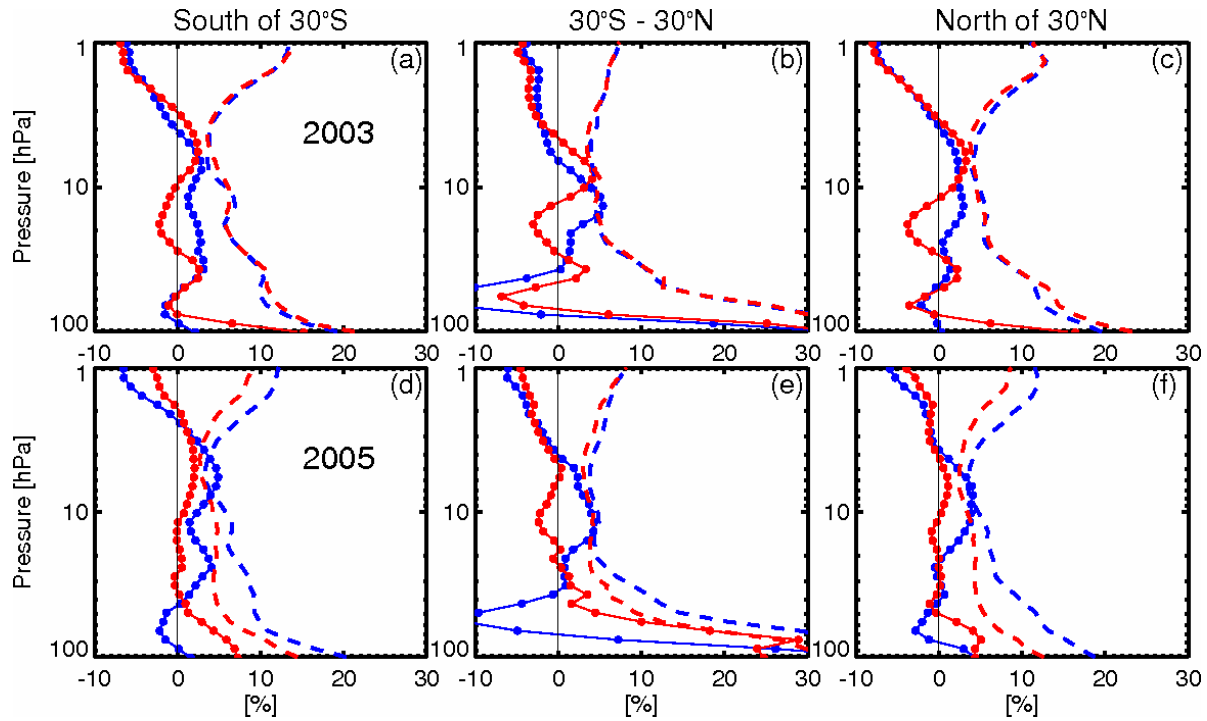
Figure 7. *Profiles of MERRA – SAGE II (blue) and MERRA-2 – SAGE II (red) comparisons relative to the SAGE II average calculated for all available data in January-August 2003 (a-c) and 2005 (d-f) and expressed in percent. Statistics for three latitude bands are shown: south of 30°S (a and d), between 30°S and 30°N (b and e) and north of 30°N (c and f). The solid lines represent the mean differences between the reanalysis and SAGE II. The standard deviations of the differences are shown as dashed lines. The SAGE II observation levels are indicated by dots superimposed on the solid lines. All statistics are shown as a percentage of the SAGE II mean over 2003 (top row) or 2005 (bottom).*
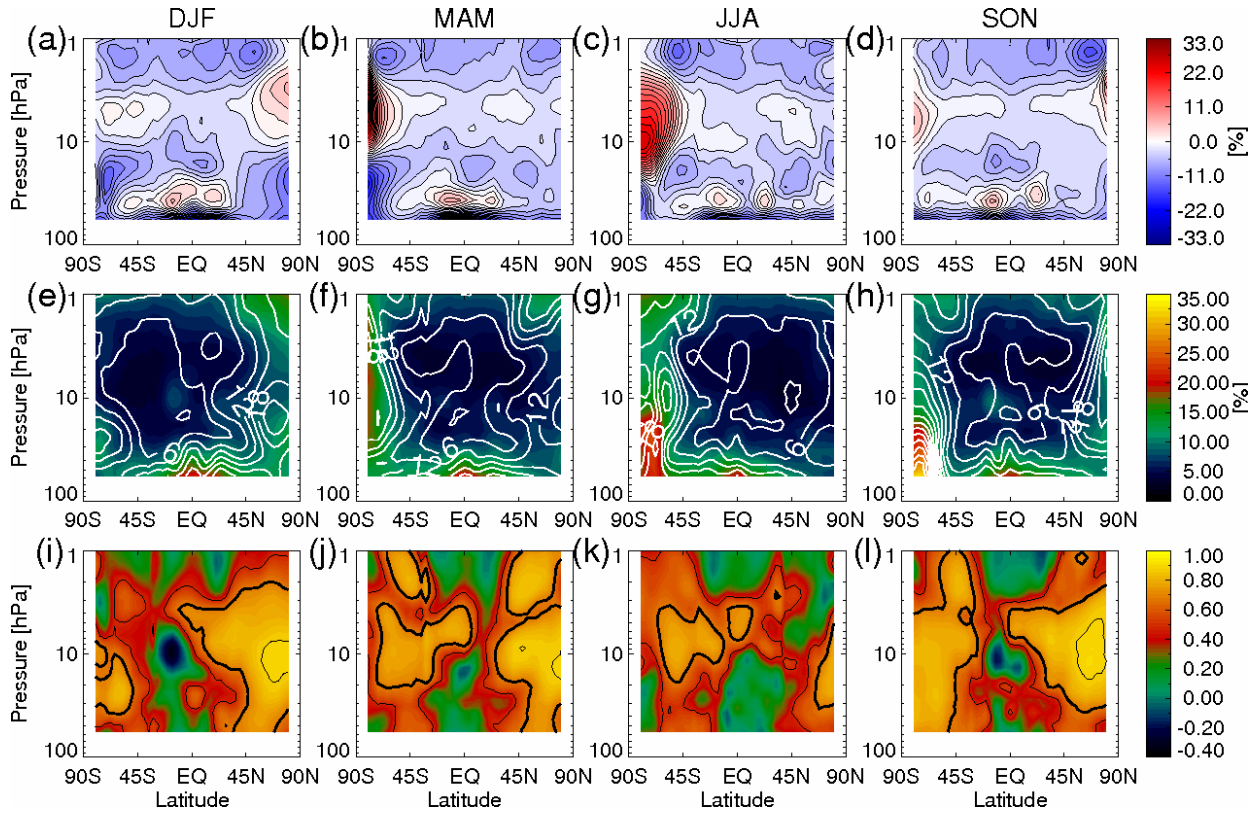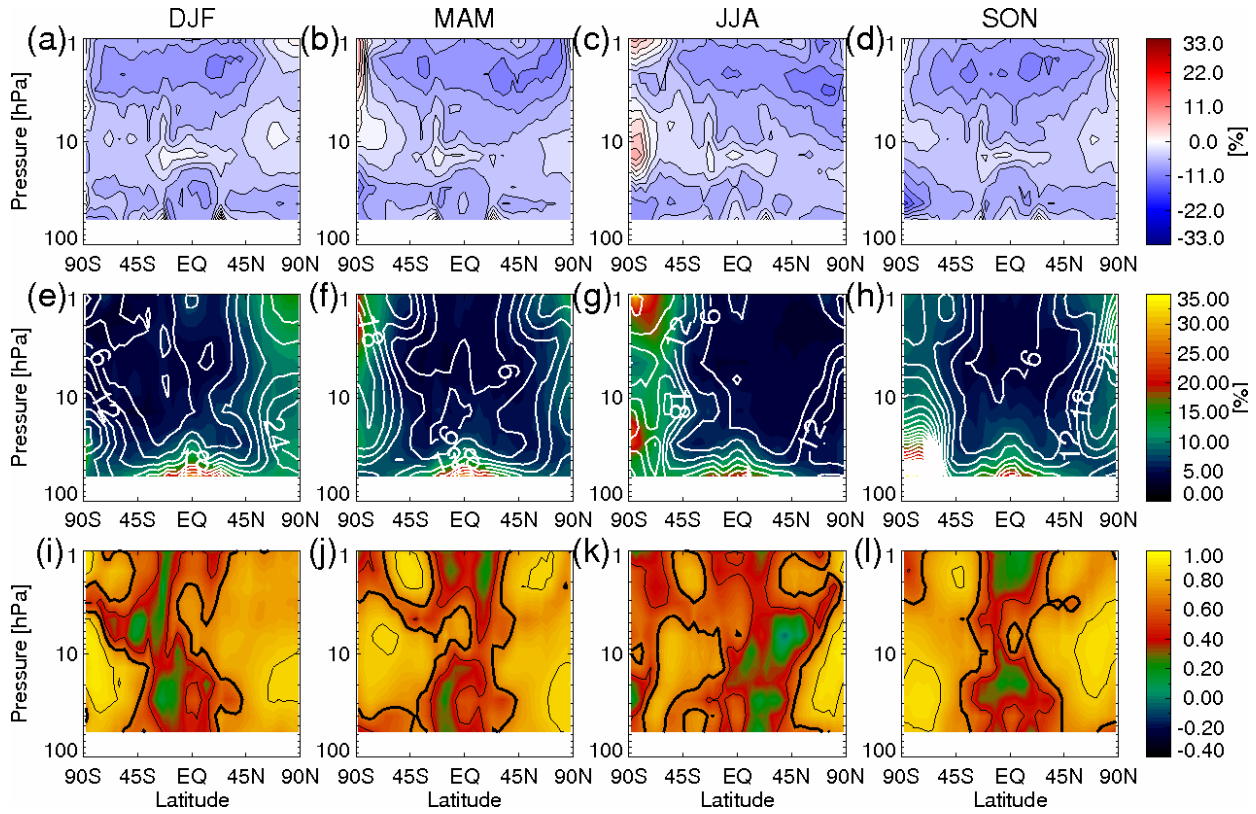
Figure 8. Top row: the zonal mean difference between MERRA-2 and UARS MLS ozone computed from all UARS MLS data between 1991 and 1996 for December – February (a), March-May (b), June-August (c) and September-November (d). The difference is expressed as a percent of the zonal mean UARS MLS ozone. Middle row: The standard deviation of the difference between the two data sets, relative to the zonal mean UARS MLS observations (colors) and relative standard deviation of UARS MLS ozone (white contours) for December – February (e), March-May (f), June-August (g) and September-November (h). Bottom row: correlations between MERRA-2 and UARS MLS for December – February (i), March-May (j), June-August (k) and September-November (l). The black contours are 0.5, 0.7 and 0.9, the 0.7 contour is thickened. The comparisons are done in the 57 hPa – 1 hPa range.
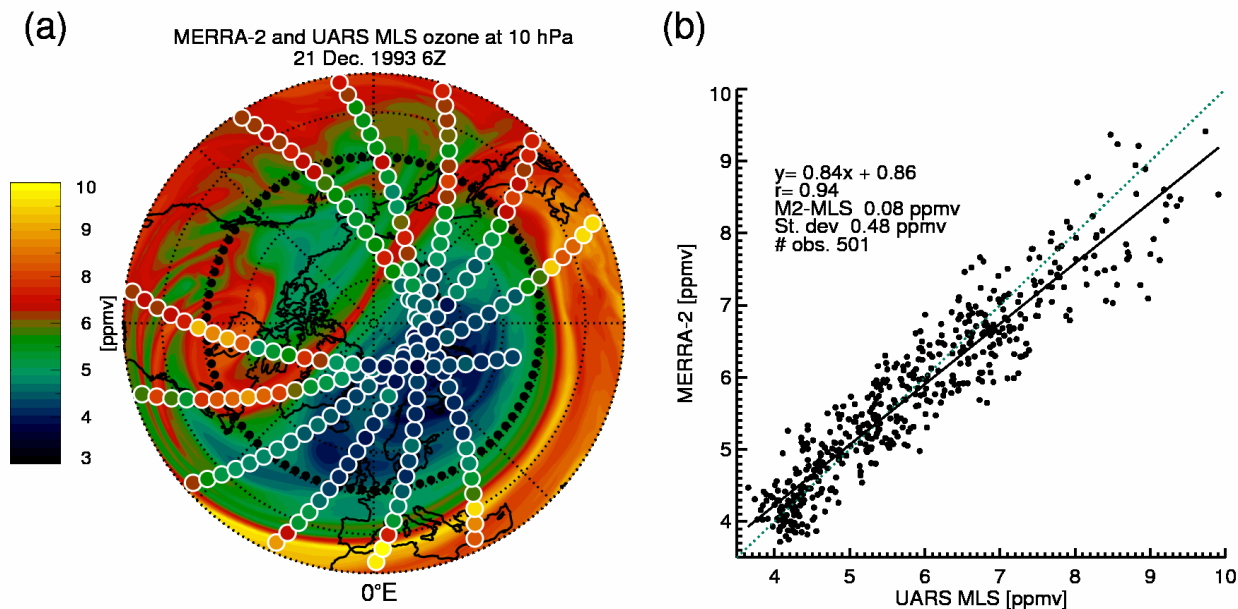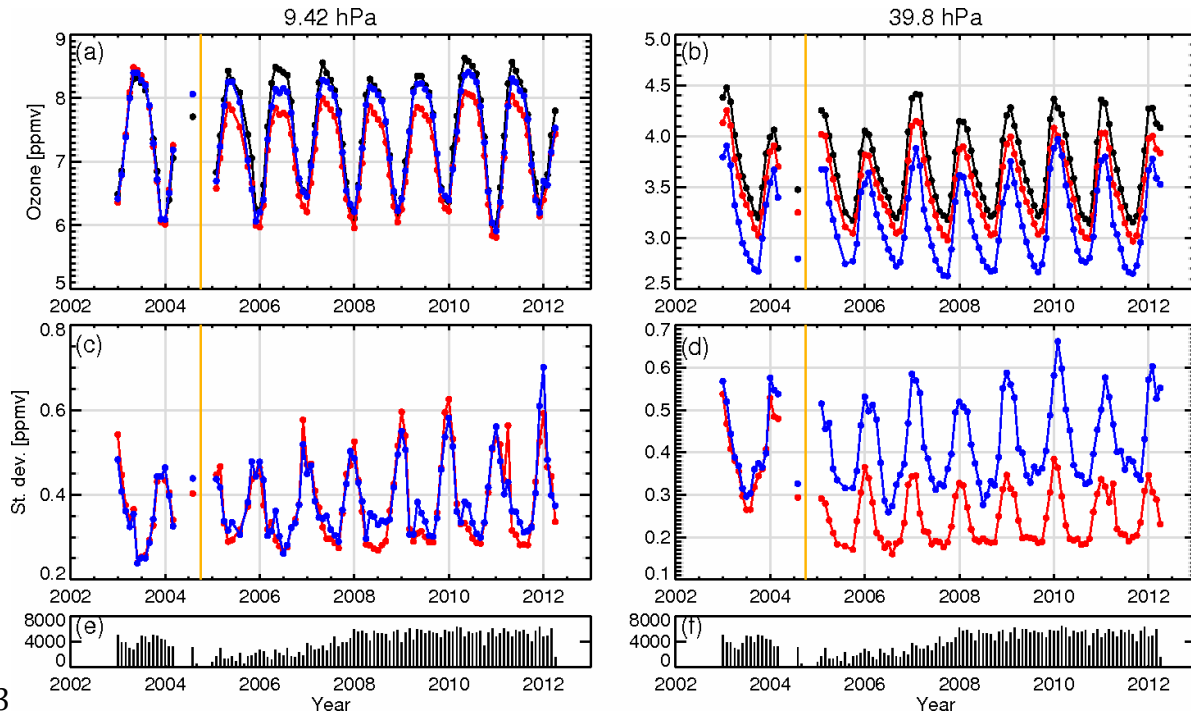
Figure 9. *(a) Color background: MERRA-2 ozone between 30°N and 90°N at 10 hPa on 21 December 1993 at 6Z. UARS MLS observations made between 1:30Z and 10:30Z are shown as filled circles color-coded by the retrieved mixing ratios. The color contour spacing is 0.2 ppmv. The highest latitude observed by SBUV around the same day (51°N) is marked by black dots. (b) a scatter plot of all UARS MLS observations at 10 hPa on 21 December and collocated MERRA-2 ozone the closest in time to the MLS data (no more than 1.5 h apart). The diagonal is shown as a dotted line and the solid black line is the linear fit. The units are parts per million by volume.*

Figure 10. *Top row: the zonal mean difference between MERRA-2 and MIPAS ozone computed from all MIPAS data between 2003 and 2012 for December – February (a), March-May (b), June-August (c) and September-November (d). The difference is expressed as a percent of the zonal mean MIPAS ozone. Middle row: The standard deviation of the difference between the two data sets, relative to the zonal mean MIPAS observations (colors) and relative standard deviation of MIPAS ozone (white contours) for December – February (e), March-May (f), June-August (g) and September-November (h). Bottom row: Correlations between MERRA-2 and MIPAS data for December – February (i), March-May (j), June-August (k) and September-November (l). The black contours are 0.5, 0.7 and 0.9, the 0.7 contour is thickened. The comparisons are done in the 57 hPa – 1 hPa range.*

Figure 11. *Time series of MIPAS, MERRA, and MERRA-2 monthly ozone statistics. 30°N – 60°N. Top row: the monthly average from MIPAS (black), MERRA (blue) and MERRA-2 (red) at 9.42 hPa (a) and 39.8 hPa (b). Middle row: the reanalysis-MIPAS difference standard deviation at 9.42 hPa (c) and 39.8 hPa (d). The units are parts per million by volume. Bottom row: The corresponding monthly MIPAS observation counts at 9.42 hPa (e) and 39.8 hPa (f). The yellow lines in panels a-d mark the transition from the SBUV to EOS Aura period in MERRA-2.*
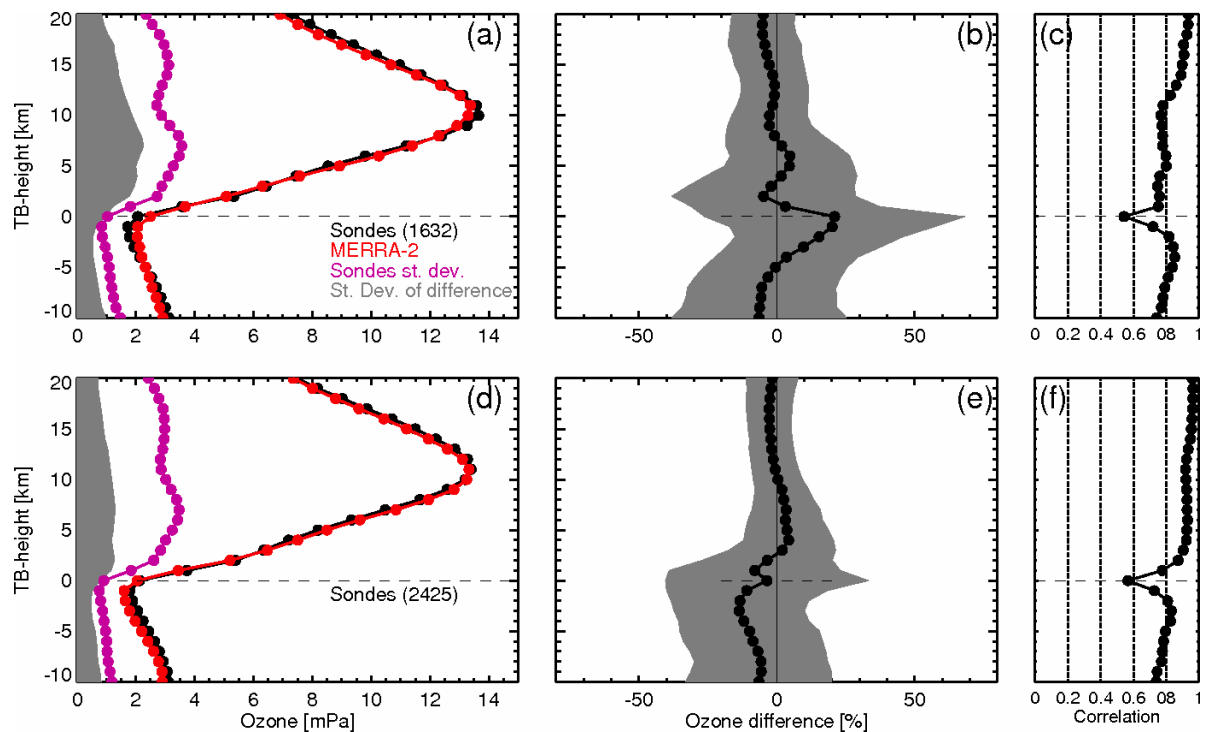
Figure 12. *Statistical comparisons of MERRA-2 ozone against ozonesondes available in 2003 (a-c) and 2005 (d-f). All data are interpolated to the tropopause-based vertical coordinate. (a and d) the average ozone partial pressure from ozone sondes (black) and MERRA-2 (red), the standard deviation of ozonesonde measurements (magenta), and the standard deviation of MERRA-2 minus ozonesonde differences (shaded). (b and e) The mean MERRA-2 minus ozonesonde differences (black line) and the standard deviation of the difference (shaded), all expressed as a percent of sonde ozone. (c and f) the MERRA-2 – ozonesonde correlations. The filled circles are the levels to which all data are interpolated and the dashed line is the dynamical tropopause. The units in a and d are millipascals.*
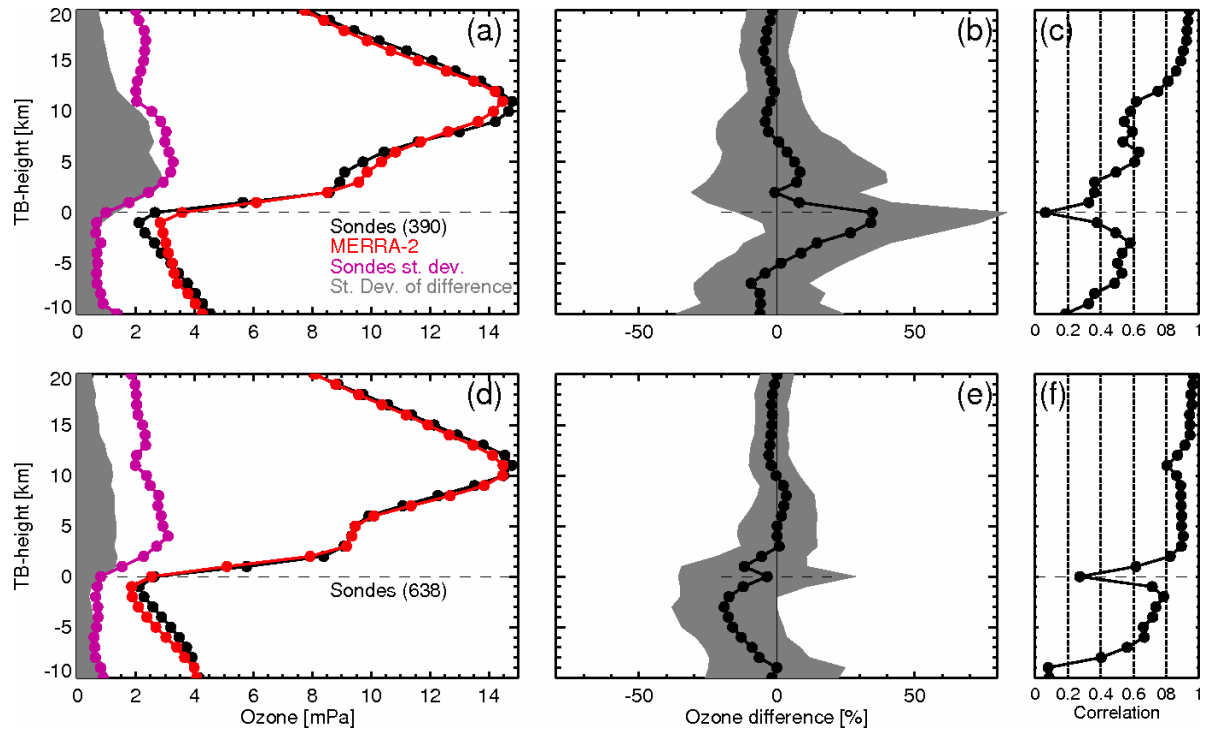
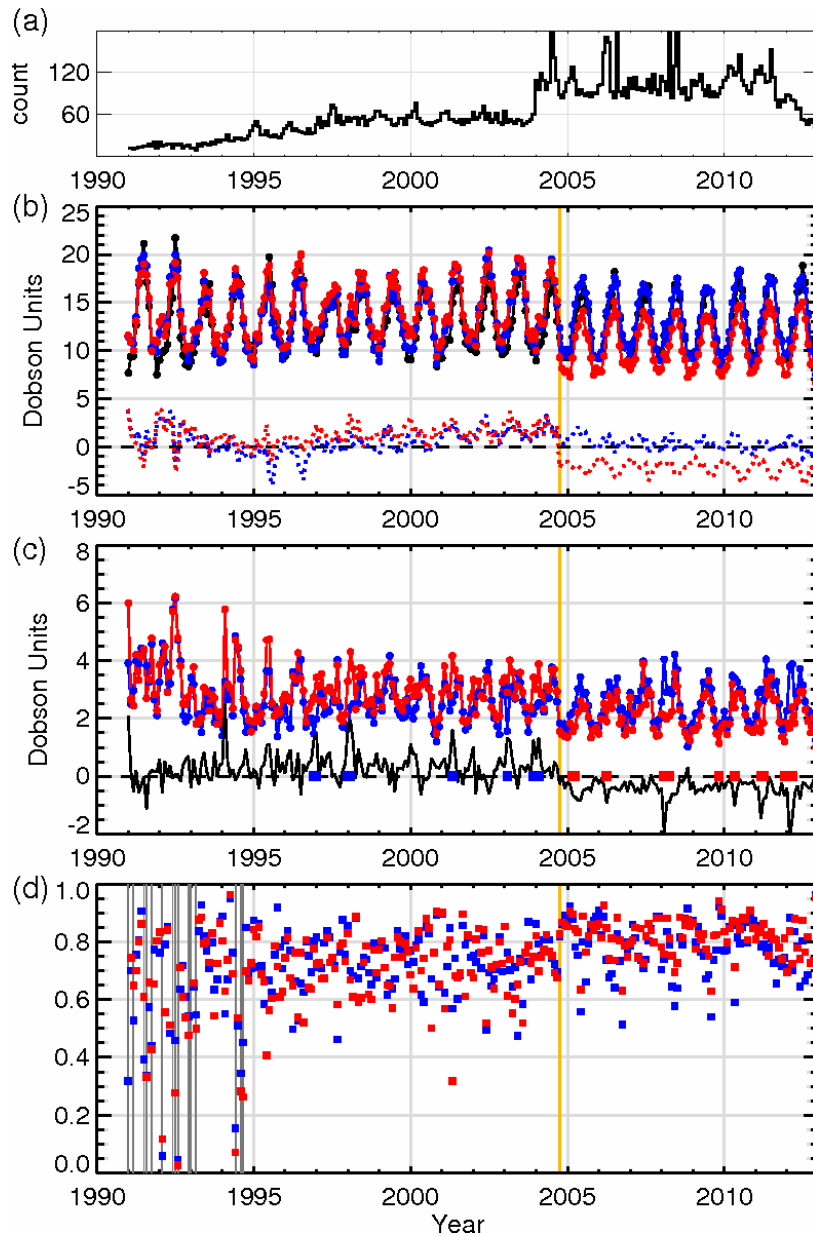**Figure 13.** *The same as Figure 13 but with soundings limited to 0°E-60°E, 45°N-60°N, March-May.*

1698
1699 **Figure 14**. *Time series of various statistics comparing ozonesondes between 30°N and*
1700 *60°N (black), MERRA (blue) and MERRA-2 (red) integrated in the upper troposphere*
1701 *(500 hPa to the dynamical tropopause). (a) number of sondes in each month. (b) the*
1702 *monthly averaged ozone (solid lines with dots) and the reanalysis minus sonde*
1703 *differences in Dobson units (dotted lines). (c) the standard deviation of the reanalysis*
1704 *minus ozonesonde differences (blue and red for MERRA and MERRA-2, respectively) and*
1705 *their difference (MERRA-2 standard deviation, minus MERRA standard deviation,, black)*
1706 *in Dobson units. The red (blue) squares on the zero line indicate months when MERRA-2*
1707 *(MERRA) is closer to the sondes than MERRA (MERRA-2) at the 95% significance level.*
1708 *(d) monthly reanalysis-ozonesonde correlations. The grey bars indicate the months for*
1709 *which the correlations are not statistically significant at 95%. The vertical yellow lines*
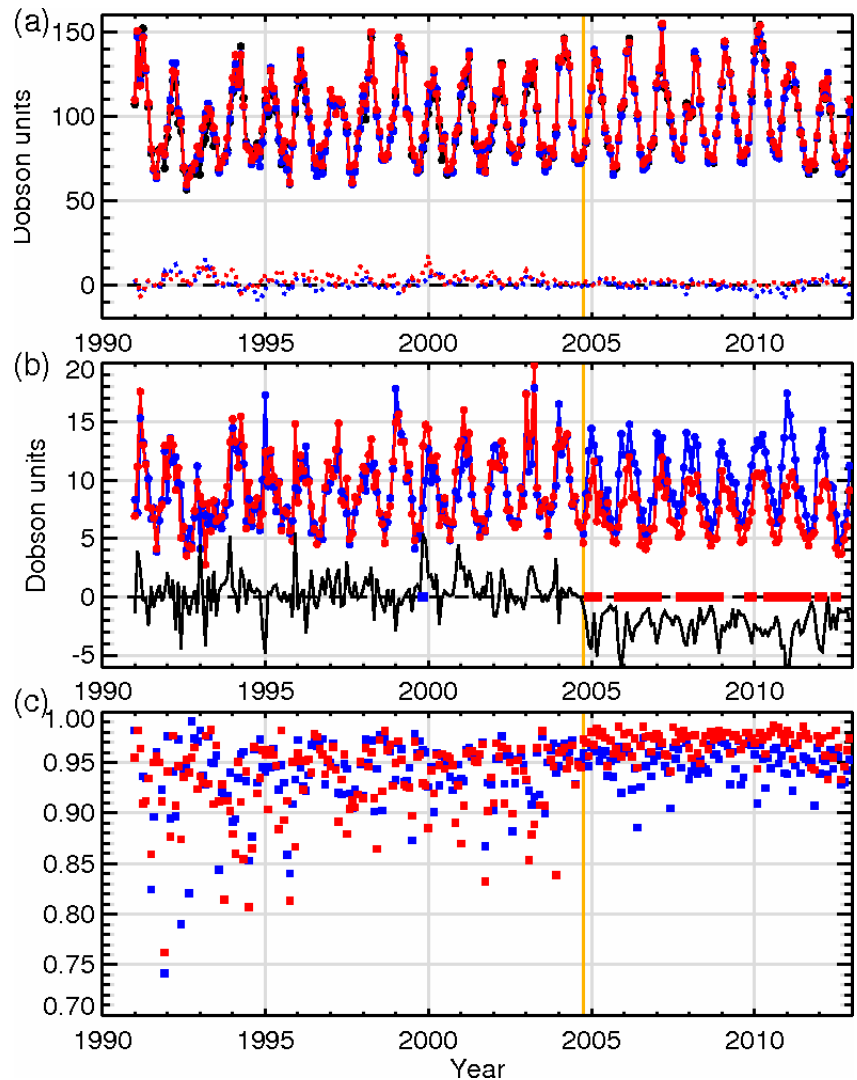1710 *indicate October 2004 when Aura data were introduced in MERRA-2.*
1711

83

1712

1713
1714 **Figure 15**. *Time series of various statistics comparing ozonesondes between 30°N and*
1715 *60°N (black), MERRA (blue) and MERRA-2 (red) integrated in the lower stratosphere*
1716 *(between the dynamical tropopause and 50 hPa). (a) the monthly averaged ozone (solid*
1717 *lines with dots) and the reanalysis minus sonde differences in Dobson units (dotted lines).*
1718 *(b) the standard deviation of the reanalysis minus ozonesonde differences (blue and red*
1719 *for MERRA and MERRA-2, respectively) and their difference (MERRA-2 standard*
1720 *deviation. minus MERRA standard deviation, black) in Dobson units. The red (blue)*
1721 *squares on the zero line indicate months when MERRA-2 (MERRA) is closer to the sondes*
1722 *than MERRA (MERRA-2) at the 95% significance level. (c) monthly reanalysis-*
1723 *ozonesonde correlations. The vertical yellow lines indicate October 2004 when Aura data*
1724 *were introduced in MERRA-2.*
1725

84

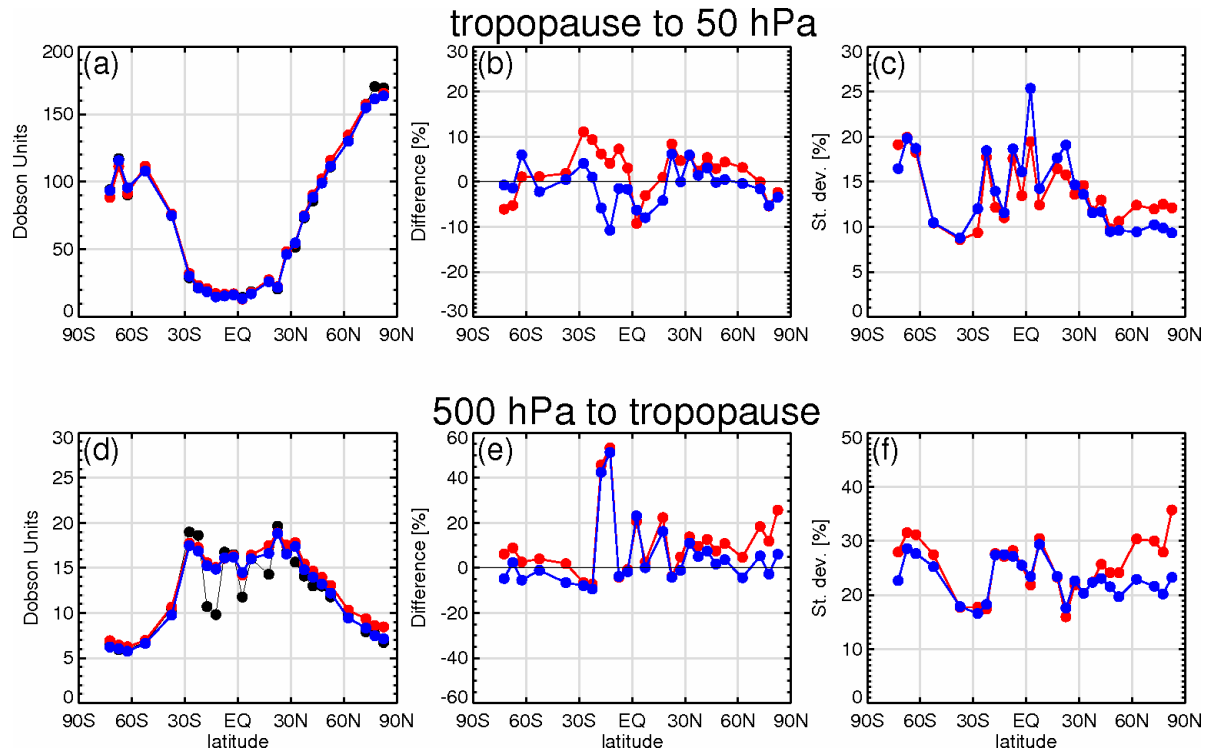## tropopause to 50 hPa

500 hPa to tropopause

Figure 16. *Comparisons of LS (a-c) and UT (d-f) ozone from MERRA (blue) and MERRA-2 (red) with ozonesondes (black) as a function of latitude. All available sondes between 1991 and 2003 are used. The data are binned into 36 latitude bins, each 5° wide. Shown are the average in Dobson units (a and d), the reanalysis minus sonde difference relative to the sonde mean in percent (b and e) and the standard deviation of the difference also relative to the sonde average in percent (c and f).*
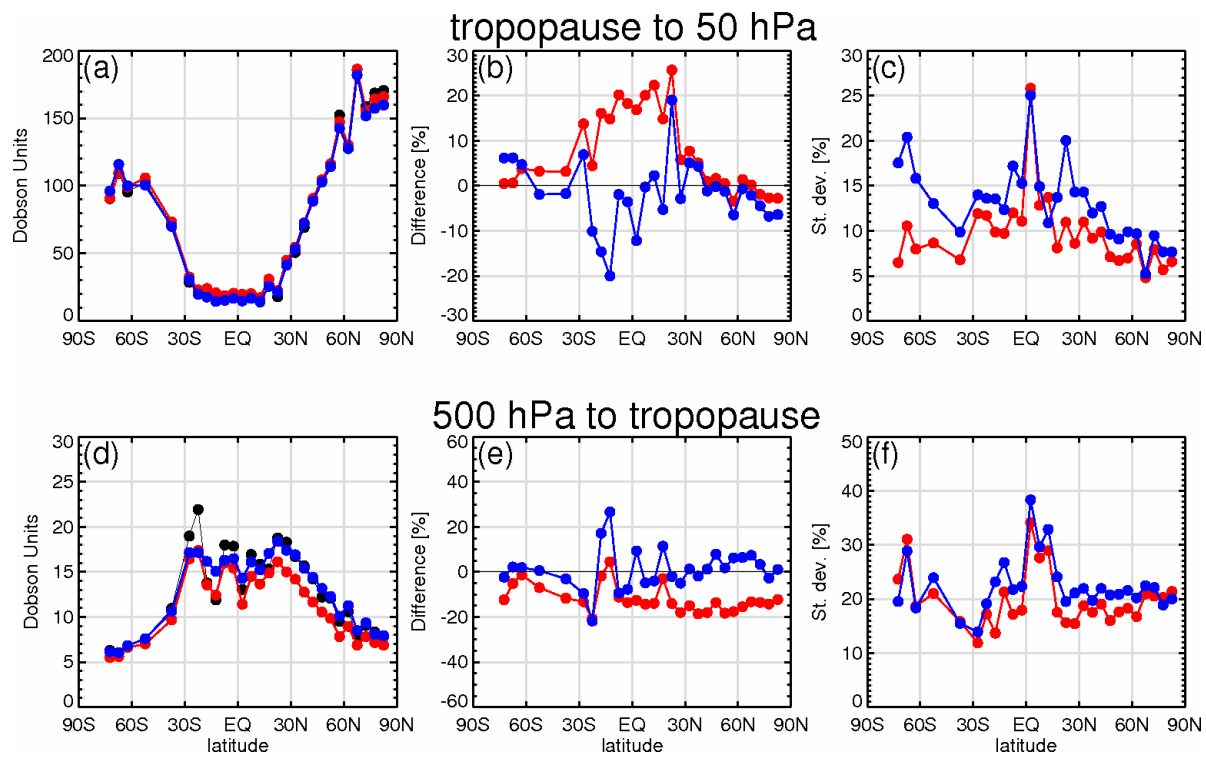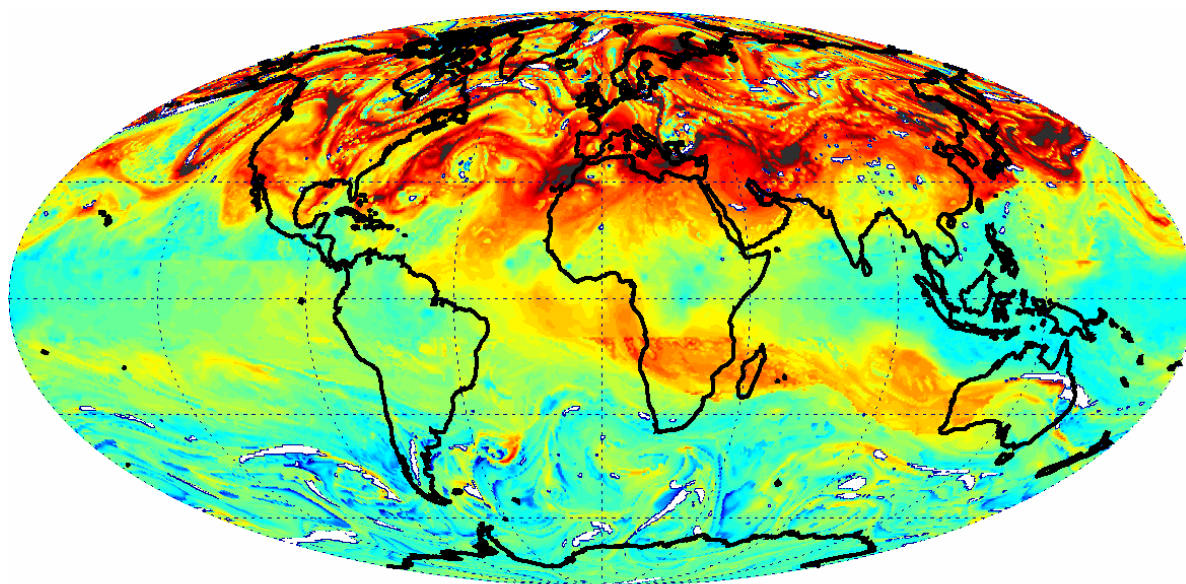
tropopause to 50 hPa

500 hPa to tropopause

Figure 17. *As in Figure 16 but for the 2005-2012 period.*

1734
1735
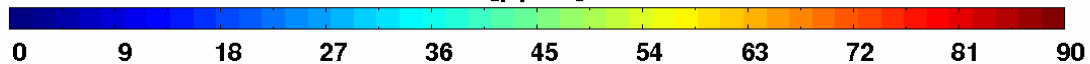1736
1737
1738
1739
1740
1741
1742
1743

Figure 18. *Mean ozone mixing ratio in parts per million by volume between 500 hPa and the tropopause from MERRA-2 on 1 July 2013. The areas where the tropopause lies below the 500-hPa surface are marked white.*